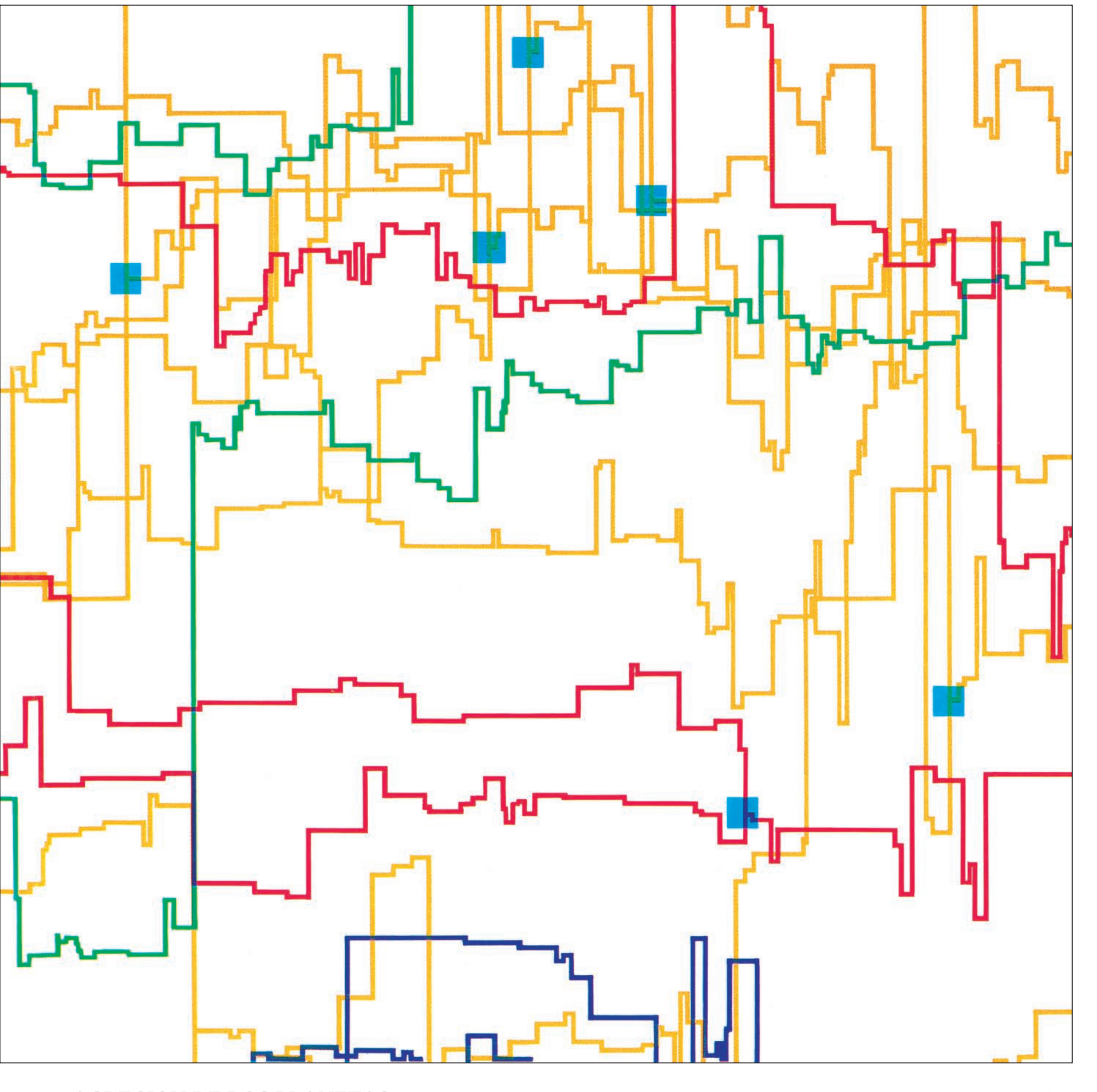
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



ACRECION DE LOS PLANETAS

Agosto 1981

250 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 59
Agosto 1981

INVESTIGACION Y CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

6 **REACTORES NUCLEARES REFRIGERADOS POR GAS, Harold M. Agnew**
Ofrecen una ventaja del máximo interés: en ellos no puede haber pérdida del refrigerante.

16 **LA DESINTEGRACION DEL PROTON, Steven Weinberg**
Las recientes investigaciones sugieren que este constituyente de la materia no es inmortal.

30 **ANTIBIOTICOS BETA-LACTAMICOS, E. P. Abraham**
Al modificar su estructura, se ha potenciado la eficacia de las penicilinas y las cefalosporinas.

48 **ARCHIBACTERIAS, Carl R. Woese**
No son procariotas ni eucariotas. ¿Nos hallamos quizás ante un nuevo reino de organismos?

62 **ASIGNACION DE RECURSOS MEDIANTE PROGRAMACION LINEAL, Robert G. Bland**
Estructuras poliédricas abstractas en espacios de muchas dimensiones ayudan a planificar.

76 **REGULACION DE LA TEMPERATURA EN EL ENJAMBRE DE ABEJAS MELIFERAS, Bernd Heinrich** El enjambre conserva el calor para volar rápidamente hacia el nuevo cobijo.

86 **FORMACION DE LA TIERRA A PARTIR DE PLANETESIMALES, George W. Wetherill**
¿Se formaron los planetas interiores del sistema solar por coalescencia de pequeños cuerpos?

98 **LOS "ICTINEOS" DE NARCIS MONTURIOL, Santiago Riera i Tuebols**
Su intuición y tenacidad resolvieron múltiples problemas de la navegación submarina.

3 AUTORES

4 HACE...

42 CIENCIA Y SOCIEDAD

110 JUEGOS MATEMATICOS

116 TALLER Y LABORATORIO

122 LIBROS

128 BIBLIOGRAFIA

EDITA: Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

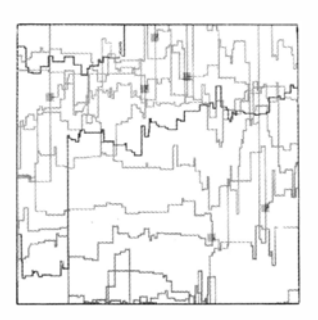
Antonio Travesí: *Reactores nucleares refrigerados por gas*; Antoni Méndez Vilaseca: *La desintegración del protón*; Cristóbal Mezquita: *Antibióticos beta-lactámicos*; Julio Rodríguez Villanueva: *Archibacterias*; Luis Bou: *Asignación de recursos mediante programación lineal*; Joandomènec Ros: *Regulación de la temperatura en el enjambre de abejas melíferas*; J. Luis Comellas García-Llera: *Formación de la Tierra a partir de planetesimales*; Luis Bou: *Juegos matemáticos*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad:

Roberto Marco y Antonio Castillo Holgado

Libros:

F. Sáez Vacas, Manuel García Velarde y Miguel Aguilar



LA PORTADA

Parte de una simulación por ordenador del proceso en el cual la colisión producida por cuerpos en órbitas elípticas alrededor del Sol llevaron a la acreción de los planetas de composición rocosa, como Mercurio, Venus, la Tierra y Marte (véase "Formación de la Tierra a partir de planetesimales", por George W. Wetherill). Cada trayectoria, de izquierda a derecha, muestra el curso de un cuerpo. La ilustración representa el semieje mayor de cada elipse, es decir, la magnitud que mide la distancia del cuerpo al Sol; los desplazamientos en vertical representan cambios en esta distancia. Los cambios están causados por colisiones y por acercamientos de unos cuerpos a otros. Cada cuadrado azul representa una colisión, de la cual surge un cuerpo mayor. Las trayectorias en amarillo representan los cuerpos que acaban por formar parte de un solo planeta, muy similar a la Tierra en su tamaño y en su posición en el sistema solar. Los otros colores representan cuerpos que se funden en tres planetas rocosos interiores.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

Condiciones de suscripción:
España:
Un año (12 números): 2.750 pesetas
Extranjero:
Un año (12 números): 43 U.S.$
Ejemplar atrasado ordinario:
280 pesetas
Ejemplar atrasado extraordinario:
420 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Miguel Munill
Balmes, 191, 2.°, 2.ª, Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Ilustración de la portada de Allen Beechel

Página	Fuente
7	George V. Kelvin
8	General Atomic Company
9-10	George V. Kelvin
11-12	General Atomic Company
13-15	George V. Kelvin
17	Ralph Morse
18	Hans W. J. Courant, Universidad de Minnesota
20-27	Gabor Kiss
31	Milton R. J. Salton, Facultad de Medicina de la Universidad de Nueva York
32-35	Ilil Arbel
36	Ilil Arbel (*arriba*); Brian G. Spratt, Universidad de Sussex
37-39	Ilil Arbel
40	Royal Society de Londres
49-54	Alan D. Iselin
55	Alexander J. B. Zehnder, Instituto Federal Suizo de Tecnología (*arriba*); Frank Mayer, Universidad de Göttingen (*abajo*)
56-60	Alan D. Iselin
63-74	Jerome Kuhl
76	Bernd Heinrich, Universidad de Vermont
78-83	Tom Prentiss
84	Tom Prentis (*arriba*); Bernd Heinrich, Universidad de Vermont (*abajo*)
86	Larry P. Cox, Instituto de Tecnología de Massachusetts
88-96	Allen Beechel
98-100	Museo Marítimo de Barcelona
101	Santiago Riera
102-103	Miguel Alonso
104	Museo Marítimo de Barcelona
105	Miguel Alonso
106	Escuela Técnica Superior de I. Industriales de Barcelona
107	Miguel Alonso
108	Deutsches Museum de Munich
110-112	Dana Nordhausen
113-115	Ilil Arbel
117	Jearl Walker
118-120	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

HAROLD M. AGNEW ("Reactores nucleares refrigerados por gas") preside la compañía General Atomic, en la que entró en 1979, después de haber dirigido durante más de ocho años el Laboratorio Científico de Los Alamos. Comenzó a trabajar en el desarrollo de la energía nuclear a principios de la década de los cuarenta, cuando recién graduado por la Universidad de Denver, se agregó al círculo restringido en torno a Enrico Fermi que trabajaba en la primera reacción de fisión nuclear en cadena. En 1943 Agnew ingresó en el Laboratorio de Los Alamos para participar en el proyecto de la bomba atómica. Desde 1946 hasta 1949 estuvo de nuevo con Fermi en la Universidad de Chicago, donde se doctoró en 1949. De allí volvió a Los Alamos. Durante tres años (1961-1964) fue asesor científico del Mando Supremo Aliado en Europa.

STEVEN WEINBERG ("La desintegración del protón") detenta la cátedra Higgins de física de la Universidad de Harvard y es científico de número del Observatorio Astrofísico Smithsoniano. Estudió en la Universidad de Cornell, por la que se graduó en 1954. Pasó un año en el Instituto Niels Bohr de Copenhague. Se recibió de doctor por la Universidad de Princeton en 1957. Desde entonces ha trabajado en la Universidad de Columbia, en el Lawrence Radiation Laboratory de la Universidad de California, en la Universidad de California en Berkeley y en el Instituto de Tecnología de Massachusetts. En 1973 entró a formar parte del claustro de Harvard. Ha escrito dos libros: *Los tres primeros minutos: Una perspectiva moderna del origen del Universo* y *Gravitación y cosmología: Principios y aplicaciones de la teoría general de la relatividad*. Ha recibido numerosas distinciones por su trabajo sobre la teoría de partículas elementales, incluyendo cinco grados honoríficos y el premio Nobel de física de 1979, que compartió con Sheldon Lee Glashow y Abdus Salam. Acaba de ser elegido miembro de la Royal Society.

E. P. ABRAHAM ("Antibióticos beta-lactámicos") es profesor emérito en la Escuela de Patología Sir William Dunn de la Universidad de Oxford. Estudió en Oxford, donde obtuvo el grado de doctor en 1938. Después de trabajar dos años en Estocolmo con una beca de la Fundación Rockefeller, regresó a Oxford para colaborar en el aislamiento y caracterización de la penicilina con Howard W. Florey, Ernst B. Chain y otros. En 1940 descubrió, con Chain, el enzima penicilinasa y, en 1953, aisló con G. G. F. Newton la cefalosporina *C* a partir de una preparación impura de penicilina *N*. Su trabajo ulterior en Oxford y en compañías farmacéuticas condujo a la introducción de las cefalosporinas en medicina. Abraham, miembro de la Royal Society, recibió el año pasado el título de Sir. En sus ratos libres gusta dar grandes paseos y esquiar.

CARL R. WOESE ("Archibacterias") enseña microbiología y genética y desarrollo en la Universidad de Illinois en Urbana-Champaign. Desde 1972 hasta 1979 explicó biofísica. En 1950 se licenció (en matemática y física) por el Amherst College, y, en 1953, se doctoró (en biofísica) por la Universidad de Yale. Pasó luegos dos años en la Facultad de Medicina y Estomatología de la Universidad de Rochester, para volver luego a Yale con el propósito de dedicarse a la investigación biofísica. Tras tres años de trabajo en la industria privada, entró en la Facultad de la Universidad de Illinois en 1964. "Toda mi carrera –escribe– ha sido una aventura hurgando en los rincones más oscuros de la evolución.". En particular, centra su interés en la historia biológica del ribosoma.

ROBERT G. BLAND ("Asignación de recursos mediante programación lineal") es profesor adjunto de la Escuela de Investigación Operativa e Ingeniería Industrial y del Centro de Matemáticas Aplicadas de la Universidad de Cornell. "Estudié en Cornell –escribe– y allí mismo alcancé el doctorado en 1974. He sido profesor ayudante de ciencias matemáticas en la Universidad estatal de Nueva York en Binghamton, becario de investigación del Centro de Investigación Operativa y Econometría de Lovaina y profesor de gerencia en el Instituto Europeo de Estudios Superiores de Dirección de Bruselas, antes de regresar a Cornell, en 1978. Mi campo de investigación son la teoría y aplicaciones de los grafos y redes, la programación matemática y los problemas de optimización discreta."

BERND HEINRICH ("La regulación de la temperatura en el enjambre de abejas melíferas") da clases de zoología en la Universidad de Vermont. "Crecí en el Maine rural, después de haber llegado a los Estados Unidos (cuando tenía 10 años de edad) procedente de Alemania. Me licencié por la Universidad de Maine y me doctoré por la de California en Los Angeles, siempre en zoología. Durante los 10 años siguientes estuve en el departamento de entomología de la Universidad de California en Berkeley. El verano pasado volví a la costa este porque añoraba las variaciones estacionales, los bosques y los ríos, que siempre han sido mi amor, mi inspiración y el origen de mi bienestar espiritual." Además de su labor científica, Heinrich es un excelente corredor de fondo, especialidad atlética en la que ha sido varias veces galardonado.

GEORGE W. WETHERILL ("La formación de la Tierra a partir de planetesimales") dirige el Departamento de Magnetismo Terrestre de la Institución Carnegie de Washington. Obtuvo sus grados académicos en la Universidad de Chicago, concluyendo en 1953 con el doctorado en física. Al finalizar sus estudios, se agregó a la plantilla del departamento que ahora preside, permaneciendo allí hasta 1960. Estuvo luego como profesor de geofísica y geología en la Universidad de California en Los Angeles. Durante cuatro de esos años (1968-72) fue jefe del departamento de Ciencia Planetaria y del Espacio. Wetherill accedió a su actual cargo en 1975.

SANTIAGO RIERA I TUEBOLS ("Los 'Ictíneos' de Narcís Monturiol") enseña historia de la ciencia y de la técnica en la Facultad de Geografía e Historia de la Universidad de Barcelona, donde se licenció. Finalizó la carrera de ingeniería industrial en 1961 y, desde 1965, ha venido profesando termodinámica y físico-química en la Escuela Superior de Ingenieros Industriales de Barcelona. Preside la sección de ingeniería de la Societat Catalana de Ciències Físiques, Químiques i Matemàtiques.

Hace...

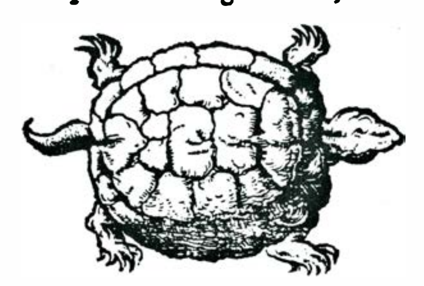
José M.ª López Piñero

... cuatrocientos cincuenta años

Andrés Laguna se trasladó a París, en cuya universidad se graduó en artes y cursó medicina. Nacido en Segovia dos décadas antes, había estudiado dos años de artes en Salamanca, pero fue en París donde se formó en idiomas clásicos junto a importantes latinistas y helenistas. Como médico tuvo por maestros a Gunther von Andernach, Jacobus Sylvius y Jean de la Ruel, es decir, los representantes más destacados del galenismo humanista parisiense. Se integró, además, en los ambientes académicos interesados por la disección y publicó, como veremos, el primer texto anatómico impreso en París.

De regreso a España en 1536, estuvo en relación con la Universidad de Alcalá, aunque no fue titular de ninguna cátedra. En opinión de Alonso Muñoyerro y de Bataillon, dio solamente un curso extraordinario en torno a 1538. Parece que asistió en Toledo a la emperatriz Isabel, pero no entró al servicio de Carlos I. Viajó a Inglaterra, vivió algún tiempo en los Países Bajos y, desde 1540 hasta 1545, residió en Metz como médico contratado por la ciudad. Durante esta última etapa pronunció en el aula magna de la Universidad de Colonia su célebre discurso pacifista sobre Europa (1543). Entre 1545 y 1554 permaneció en Italia, recibiendo en Roma diversos honores concedidos por los papas Pablo III y Julio III. También estuvo en Venecia, alojado en la casa del embajador Juan Hurtado de Mendoza, importante humanista y dueño de una de las mejores bibliotecas de la Europa de esta época. A mediados de 1554 se trasladó a los Países Bajos, donde estuvo hasta finales de 1557, fecha en la que volvió a España. Falleció dos años más tarde, seguramente en Guadalajara.

ANATOMICA METHODVS, SEV DE SECTIONE HVmani corporis contemplatio, ANDREA A LACVNA Secobiēſe authore, in cōpendium atq; adeò enchiridium redacta: vbi quicquid boni ſucculētive probatiſſimi quique philoſophorum de ea ſcripſerunt, appoſitiſſime congeſtū reperias: obſeruata quidem ſermonis puritate, at verborum lenociniis, ac, vt ita dicam, phaleris, nō admodum anxie curatis: nam & quę latinæ voces plerūq; nō ſup peterent, græcis vti præſtitit, q̄ rem in medio relinquere. Eme igitur lector, ac fruere



TECVM HABITA!

Pariſiis apud Ludouicum Cyaneum.

M. D. XXXV.

Portada de **Anatomica methodus** *(1535), de Andrés Laguna, primer texto morfológico impreso en París*

Laguna publicó más de treinta obras, algunas de las cuales son de tema literario, histórico, filosófico y político. Bataillon ha aducido razones de peso para considerarlo autor del famoso *Viaje de Turquía,* aunque esta adscripción ha sido criticada por Dubler y otros investigadores. Tradujo y comentó los ocho últimos libros de la agronomía atribuida a Constantino IV Pogonatos (1541), así como los textos aristotélicos *De physiognomicis* (1535), *De mundo* (1538) y *De natura styrpium* (1543). No obstante, la mayor parte de su producción científica –y la más importante– corresponde a la medicina.

Como hemos adelantado, Laguna escribió un texto morfológico, que tituló *Anatomica methodus* (1535), cuando realizaba estudios en París. Bajo el influjo de las traducciones del griego al latín de las obras anatómicas de Galeno, la capital francesa se había convertido en el centro de la anatomía galénica pura. La enseñanza académica continuaba dándose a la usanza tradicional, pero al margen de la misma alcanzó gran importancia la práctica de disecciones privadas. Laguna fue, solo o en colaboración con Jean Tagault, uno de los primeros que realizó dichas disecciones. La estructura general de su libro y la mayor parte de su contenido son estrictamente galénicos, lo que no impide que incluya los resultados de sus observaciones en cadáveres humanos. Entre ellos se encuentran algunos de cierta importancia, como la descripción de un riñón en herradura o la comprobación de la existencia de dos orificios en el ciego. Este último hallazgo había sido ya expuesto por Mondino de Luzzi, al que Laguna cita expresamente, y también por Berengario de Carpi. No obstante, la contribución de Laguna es muy anterior a la de Gaspar Bauhin, Guillaume Rondelet y otros autores, tan arbitrariamente considerados "descubridores de la válvula ileocecal", como Laguna lo fue por sus panegiristas españoles. La *Anatomica methodus* no debe, sin embargo, su relieve histórico a la importancia de sus hallazgos, sino a la energía con la que criticó la forma tradicional de enseñar anatomía, que dejaba en manos de un barbero la práctica de la disección. Frente a ella, el médico español subrayó la necesidad de realizarla por sí mismo y de atenerse con rigor a los hechos cuidadosamente observados. Este es el aspecto que ha destacado O'Malley al hablar de la influencia de Laguna sobre Vesalio, contemporáneo suyo algo más joven que coincidió con él en París. Ello resulta coherente con el respeto que más tarde manifestó siempre Vesalio por el médico segoviano. Elogió su labor de comentarista de Galeno y su consejo fue decisivo para que aceptara el tratamiento "suave" de las heridas por arma de fuego, una de las novedades más significativas del empirismo quirúrgico de la época.

Las publicaciones propiamente médicas de Laguna se centraron en la obra de Galeno desde los supuestos de la corriente humanística. Aunque tradujo al latín dos escritos galénicos, tras depurar el texto griego, su principal aportación en este sentido fue el *Epitomes omnium Galeni Pergameni operum* (1548). Se trata de una exposición en tres volúmenes de las doctrinas del médico griego, resumiendo sus obras a partir de las traducciones latinas de Erasmo de Rotterdam, Thomas Linacre, Gunther von Andernach, Niccolo Leoniceno, etcétera. Fue muy apreciada en toda Europa, siendo reeditada siete veces. En una carta dirigida a su cuñado Martinus Astriacus que figura al final de esta obra, Vesalio hizo un expresivo elogio de Laguna: "Vir graecis latinisque iuxta ac pari lege peritus". El libro incluye una *Vita Galeni* y fue ampliado notablemente por su autor en la reedición de 1551. Desde entonces se reimprimió unida a unas *Annotationes in Galeni interpretes* (1548), en las que Laguna critica los errores que, en su

opinión, habían cometido los traductores y comentaristas anteriores. La transición al galenismo "hipocratista", que culminaría más tarde en figuras como Francisco Valles, se refleja en el libro que dedicó Laguna a los comentarios de Galeno a Hipócrates (1554), que lleva como apéndice un tratadito titulado *Nonulla Galeni Enantiomata,* dedicado a glosar una veintena de contradicciones del médico de Pérgamo.

Laguna publicó, por otra parte, estudios sobre la peste (1542 y 1556) y sobre la gota (1551), tratados de dietética (1546 y 1547), un breve texto acerca de los pesos y medidas farmacéuticos y una monografía consagrada a las carúnculas uretrales (1551). Riera ha demostrado que esta última es la primera contribución al tema en la literatura científica europea.

La obra más conocida del médico segoviano es, sin embargo, su traducción castellana con comentarios de la *Materia medica* de Dioscórides. No solamente se preocupó de establecer un texto griego lo más depurado posible, sino que comprobó las descripciones de Dioscórides, herborizando en numerosos lugares de Europa y en distintas zonas del Mediterráneo. La traducción es clara y precisa y los comentarios constituyen una riquísima fuente, no sólo para la botánica médica de la época, sino para otras muchas actividades científicas y técnicas. Fue, además, uno de los textos científicos españoles de esta época de más prolongada pervivencia. Impresa por vez primera en Amberes el año 1555, se reeditó en veintidós ocasiones en España hasta finales del siglo XVIII. Conviene advertir, sin embargo, que su importancia en el contexto científico español de su tiempo ha resultado exagerada, en parte debido a dicha pervivencia y quizá también porque es uno de los textos mejor atendidos por la investigación histórica. Sus limitaciones son, en efecto, tan innegables como sus cualidades. Laguna se basó en buena parte en la traducción del libro de Dioscórides que había publicado el italiano Piero Andrea Mattioli (1544) y las ilustraciones que incluye proceden en casi su totalidad de la obra de este último. Por otra parte, ni por la novedad de sus materiales, ni por el detalle de sus descripciones, puede esta obra equipararse a los estudios sobre el Nuevo Mundo de los naturalistas españoles de la época.



La encina y el alcornoque. Grabados de la traducción castellana con comentarios de la **Materia medica** *de Dioscórides (1555), la más conocida de las obras de Andrés Laguna*

Reactores nucleares refrigerados por gas

Empleados con éxito en varios países europeos, poseen una característica muy atractiva en lo concerniente a la seguridad: en ellos, resulta imposible un accidente de pérdida del refrigerante, como ocurrió en Three Mile Island

Harold Agnew

En marzo de 1979 la industria de la energía núcleo-eléctrica sufrió un shock del que todavía no se ha recuperado: el accidente que inutilizó uno de los reactores de Three Mile Island. Por ironía de las cosas, un suceso que no lesionó a nadie arruinó las perspectivas de expansión de la energía nuclear, precisamente en el mismo momento en que el pueblo norteamericano estaba empezando a tomar conciencia de la necesidad de nuevas fuentes nacionales de energía.

Aunque la experiencia de Three Mile Island les demostró, de forma muy satisfactoria, a las personas técnicamente calificadas que los actuales reactores nucleares no presentan serias amenazas contra la salud y seguridad del público, reveló también que esta clase de accidentes y fallos en los equipos pueden poner en peligro la operabilidad de la planta y arriesgar las grandes inversiones de capital que una central de este tipo supone. Y en una situación límite, un accidente como el de Three Mile Island puede hacer tambalear la supervivencia financiera de la compañía propietaria de la instalación.

Tal vez la principal lección que se desprende de lo ocurrido en Three Mile Island sea que las actuales plantas nucleares para la producción de energía son vulnerables a muy raros sucesos que pueden darse en una situación en que el tiempo disponible para responder correctamente no llegue a un minuto. En tales casos, aunque muy poco probables, si las acciones correctoras apropiadas no se ponen en práctica de inmediato, las consecuencias podrían ser costosísimas, por más que no exista el problema de la seguridad del público. Resulta, pues, razonable preguntarse: ¿podemos quedar tranquilos con reactores nucleares diseñados de tal modo que los operadores, para evitar que se dañe la planta, deben reaccionar correctamente en menos de un minuto? No cabe respuesta afirmativa.

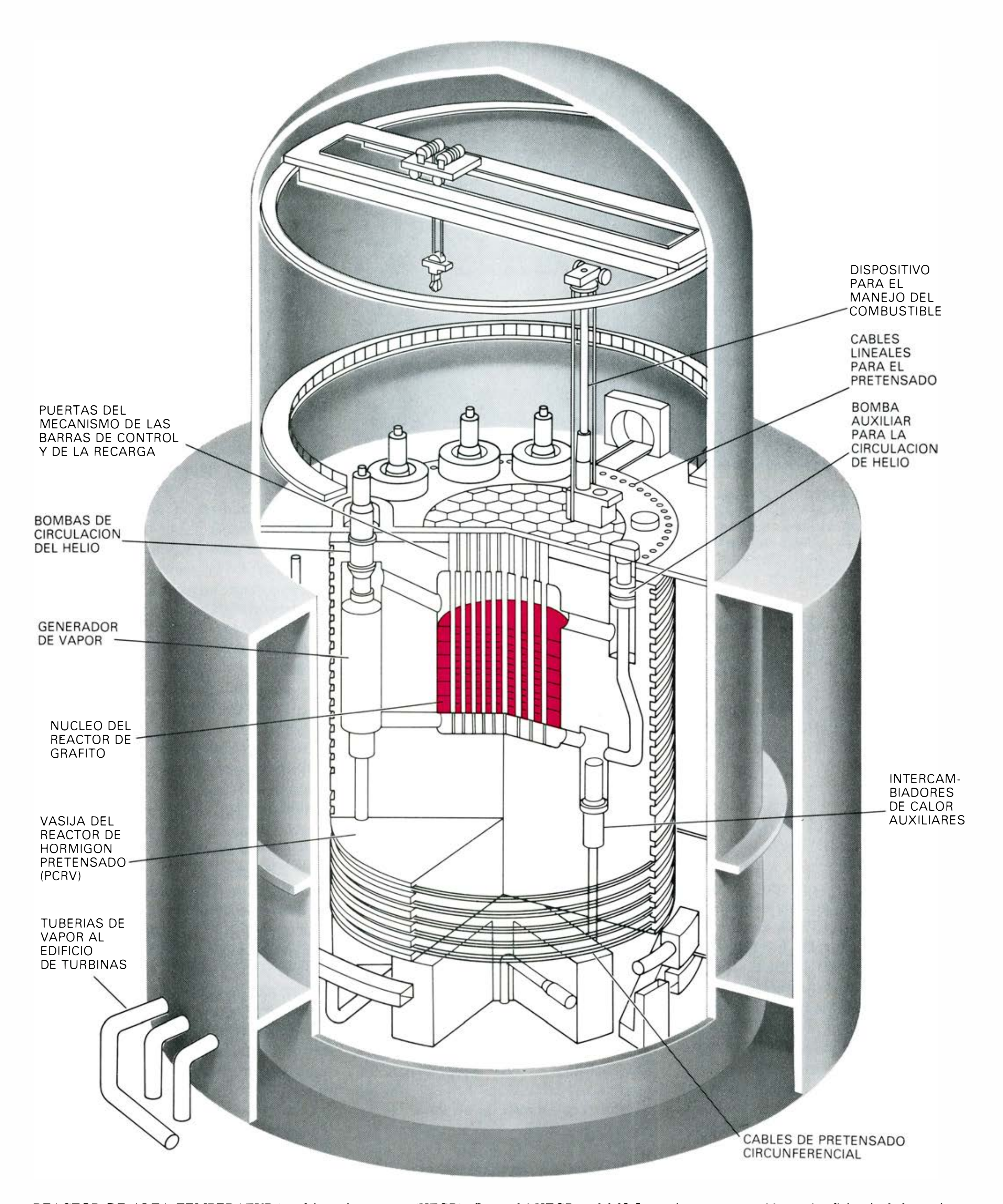
Si tal es el planteamiento, ¿por qué la industria núcleo-eléctrica de los Estados Unidos ha seguido esa política? El predominio del reactor de agua ligera en Norteamérica tiene una explicación muy sencilla. El reactor de agua ligera a presión constituye una adaptación directa del reactor, altamente compacto, diseñado para propulsar el primer submarino nuclear: el U.S.S. *Nautilus*, botado en 1954. Una versión del reactor del submarino para la producción de energía eléctrica construido por la Westinghouse Electric Company entró en servicio en Shippingport, Pennsylvania, tres años después. Al poco, la General Electric Company introdujo un reactor de su propio diseño, el reactor de agua en ebullición, donde el calor generado por la fisión nuclear se extraía del núcleo por vapor, no por agua caliente a presión.

En ambos tipos de reactor es esencial que el núcleo del reactor no llegue a estar sin cubrir por el agua, aunque sea un tiempo muy breve, para que la temperatura del núcleo no suba rápidamente y funda las vainas metálicas que encierran las pastillas de combustible, como probablemente ocurrió en Three Mile Island. Los reactores de agua ligera están equipados con mecanismos de seguridad redundantes para evitar accidentes por "pérdida de refrigerante". En tales accidentes, el equipo de emergencia está proyectado para inundar el núcleo con agua desde una fuente abundante y segura. Cuando el flujo normal de refrigeración se interrumpió en Three Mile Island, una cadena de sucesos muy improbables, a los que hay que sumar, según todos los indicios, un error del operador, interrumpió la entrada de la cascada de agua de emergencia en el núcleo durante un tiempo demasiado largo.

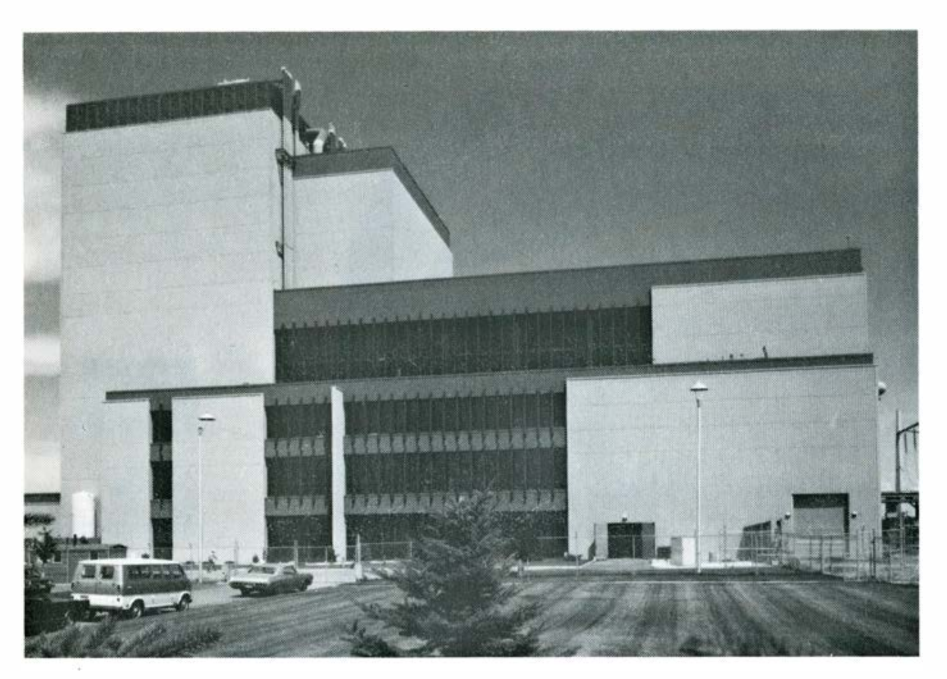
Salvo uno, los 71 reactores de las centrales nucleares autorizadas y en funcionamiento en los Estados Unidos, que suministran cerca del 11 por ciento de la energía eléctrica consumida por la nación, son reactores de agua ligera. La excepción es un reactor refrigerado por helio, en la Central Nuclear de Fort St. Vrain, puesto en servicio en el verano de 1979 por la Compañía de Servicios Públicos de Colorado. La potencia nominal de la planta es de 330 megawatt de potencia eléctrica o MWe, que viene a ser un tercio de la potencia de una central comercial típica. El reactor ha estado operando al 70 por ciento de su potencia nominal. La Comisión Reguladora Nuclear acaba de conceder autorización para rendir a plena potencia.

La central piloto de Fort St. Vrain, incluida en el Programa de Prueba de Reactores de Potencia de la Comisión de Energía Atómica, fue diseñada y construida, para la Compañía de Servicios Públicos de Colorado, por la General Atomic Company. Le precedió el funcionamiento satisfactorio de un prototipo de 40 MWe, la Central Nuclear de Peach Bottom No. 1, en la red de producción de la Compañía Eléctrica de Filadelfia. En sus siete años y medio de actividad, de 1967 a 1974, el reactor de Peach Bottom se mantuvo útil para el servicio el 86 por ciento del tiempo (sin contar las paradas programadas para investigación y desarrollo del propio reactor). Ese parámetro se cifra en el 66 por ciento en todos los reactores nucleares norteamericanos.

Por razón de la seguridad, los reactores refrigerados por helio de distinguen de los refrigerados por agua en dos rasgos fundamentales. Primero, dado que el núcleo del reactor está refrigerado



REACTOR DE ALTA TEMPERATURA refrigerado por gas (HTGR). Se trata de un sistema correspondiente a la segunda generación, más eficiente que los 71 reactores de potencia que suministran hoy cerca del 11 por ciento de la electricidad de los Estados Unidos. En este HTGR diseñado por la General Atomic Company, el moderador (el material que frena los neutrones en el núcleo del reactor) es grafito y, el refrigerante, helio. En los reactores de agua ligera normales se utiliza agua (desmineralizada y acondicionada) como moderador y refrigerante. El HTGR que se indica aquí produciría 860 megawatt (MWe) de electricidad, cifra ligeramente inferior que la de las mayores centrales nucleares, que generan 1000 MWe o más. El rendimiento térmico del HTGR es del 38,5 por ciento, comparable con la eficiencia de las mejores plantas que utilizan fuel de combustible, y es más alta que el 32 o 33 por ciento obtenido por los actuales reactores de agua ligera. Como el núcleo del HTGR contiene casi 1500 toneladas de grafito, dotado de una gran capacidad de absorción de calor, la probabilidad de que se produzcan en el núcleo daños irreparables si hay una interrupción en el flujo del refrigerante o una pérdida del mismo es muy inferior en el HTGR que en el reactor de agua ligera. Una interrupción de este tipo causó el accidente de Harrisburg. El núcleo del reactor y el sistema de generación de vapor están alojados en una vasija de hormigón pretensado, cuyas paredes miden unos 460 centímetros de grosor.



EL UNICO REACTOR HTGR que poseen los Estados Unidos es la Estación Generadora de Fort St. Vrain, proyectada por la General Atomic y operada por su propietaria, la Compañía de Servicios Públicos de Colorado. La planta, que tiene una potencia de 330 MWe, entró en servicio hace tres años y medio; de entonces acá ha producido más de dos mil millones de kilowatt-hora de electricidad. En varias ocasiones se ha interrumpido la circulación forzada del refrigerante en el núcleo a lo largo de intervalos de hasta 15 minutos de duración, sin producirse ningún daño en el núcleo o en los elementos combustibles.

por un gas que circula enteramente confinado en el interior de la vasija, no cabe pensar en pérdidas de refrigerante primario por rotura de tuberías fuera de la vasija. Segundo, si la circulación del gas se interrumpe por algún malfuncionamiento del sistema principal de circulación del helio, la temperatura del interior del núcleo del reactor sube muy lentamente, porque los elementos combustibles están embebidos en el interior de una masiva matriz de grafito, que actúa de moderador para el frenado de los neutrones y puede absorber el calor desprendido por los productos de fisión, una vez parada la reacción en cadena. En un reactor moderado con grafito y refrigerado con helio, la reacción nuclear se detiene mediante la inserción de barras de control, como en los demás reactores, o bien mediante la inyección de pequeñas bolas, conteniendo boro que "envenena" la reacción. En los reactores refrigerados por agua, la pérdida del refrigerante, que también actúa como moderador, suspende la reacción.

Si en un reactor refrigerado por agua fallara el funcionamiento del sistema de refrigeración de emergencia, se elevaría la temperatura del núcleo del reactor, incluso después de detenerse la reacción, debido a que los productos de fisión acumulados en los elementos combustibles continuarían desprendiendo energía en gran cantidad. En el momento de la parada, el calor residual de la desintegración de los productos de fisión equivale al 7 por ciento de la potencia térmica nominal, o sea, alrededor de 210 megawatt en una planta refrigerada por agua de 3000 megawatt de potencia (equivalente a una potencia eléctrica de 1000 MW). Se ha calculado que, en un accidente de pérdida de refrigerante de este tipo, la temperatura de la vaina de los elementos combustibles alcanzaría unos 1650 grados Celsius y que los fallos en los elementos combustibles tardarían unos 50 segundos en producirse en un reactor de agua a presión y menos de dos minutos en un reactor de agua hirviendo. Si ocurriera algo parecido en un reactor refrigerado por helio, con una despresurización del conjunto y fallo total del sistema de circulación de helio, se requiriría más de una hora para que la temperatura del interior del núcleo alcance los 1650 grados Celsius. A esta temperatura, no quedarían afectados ni las partículas de combustible dentro de su vaina ni los elementos combustibles de grafito. Las partículas de combustible y el grafito pueden resistir fácilmente temperaturas de hasta 2200 grados Celsius, que tardarían en formarse, además, 10 horas por lo menos. En resumen, hay un amplio margen de tiempo para programar múltiples medidas de emergencia relativas a la restauración del flujo del refrigerante de helio.

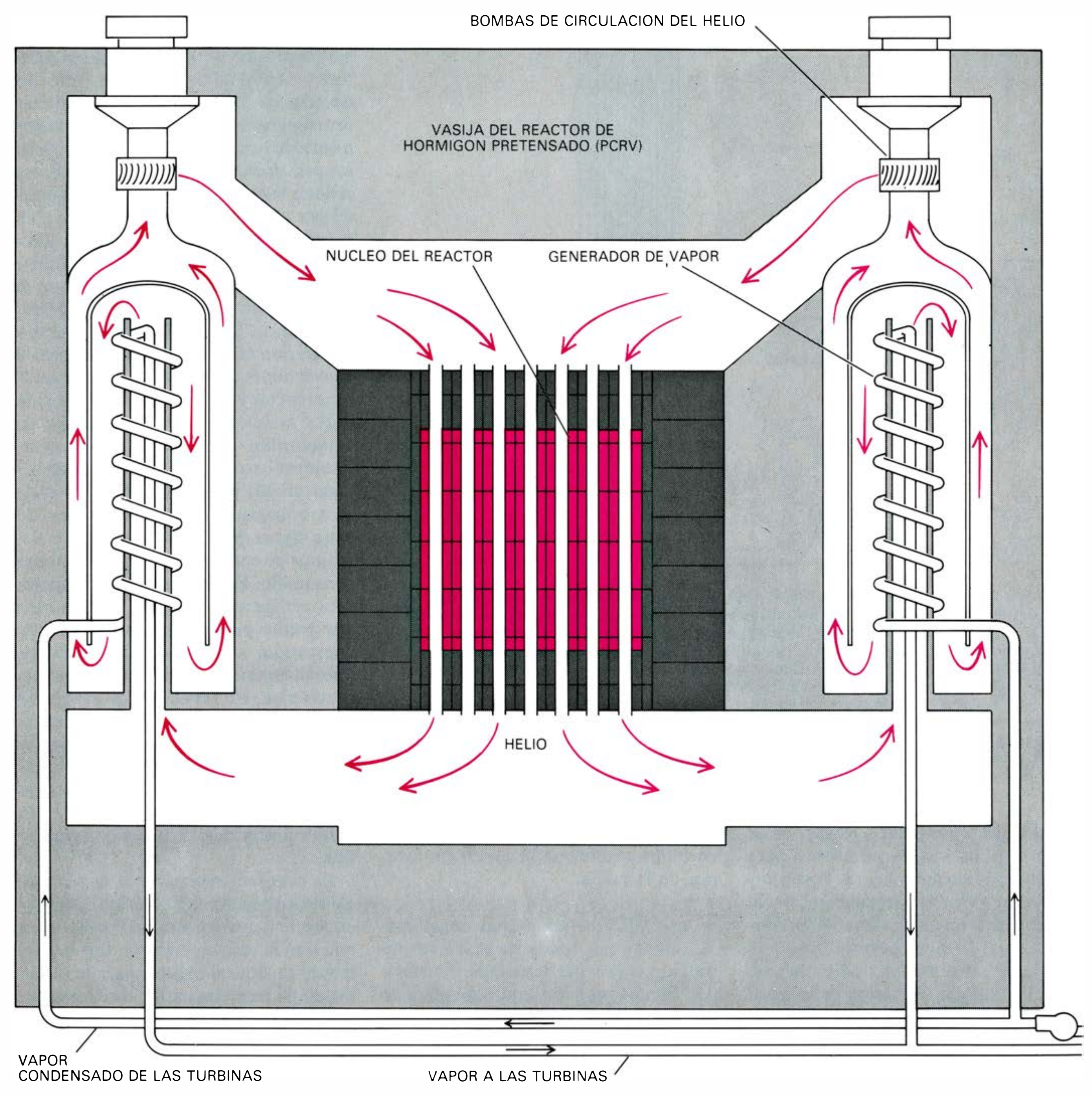
Las virtudes de los reactores de grafito refrigerados por gas gozan de amplio reconocimiento. En las décadas de 1950 y 1960, cuando los Estados Unidos estaban empeñados en los reactores de agua ligera, Gran Bretaña y Francia desarrollaron los reactores moderados con grafito y refrigerados con gas, en los que el refrigerante era dióxido de carbono, en vez de helio. Gran Bretaña dispone hoy de más de 40 reactores refrigerados por gas, en servicio o en fase de construcción. Francia tiene siete, e Italia, España y Japón uno por nación. Con los reactores europeos refrigerados por gas se ha adquirido más de 600 años-reactor de experiencia de operación. Estos reactores han producido casi un quinto de la energía eléctrica de generación nuclear de Europa, Japón y Estados Unidos.

El interés británico y francés arranca de 1956, cuando un grupo de físicos, muchos de los cuales habían trabajado en el Laboratorio Científico de Los Alamos, se reunieron en La Jolla, California, para examinar el problema del diseño de un reactor que obtuviera un rendimiento mayor y requiriera menos atenciones que los reactores entonces disponibles. Entre los asistentes se encontraban H. A. Bethe, de la Universidad de Cornell, Freeman J. Dyson, del Instituto de Estudios Avanzados, Peter Fortescue, del Instituto de Investigación de Energía Atómica de Harwell en Inglaterra, y Frederic de Hoffmann, presidente a la sazón de la General Atomic. De estas primeras conversaciones, apoyadas por la opinión de británicos y franceses, surgió la idea del reactor de alta temperatura refrigerado por gas, o HTGR ("*H*igh-*T*emperature *G*as-cooled *R*eactor"), ensayado en Peach Bottom y, a mayor escala, en Fort St. Vrain. Como Estados Unidos suministraba helio en abundancia, pudo seleccionarse este gas de refrigerante, en vez de CO_2. El helio tiene la ventaja de su estabilidad a altos flujos de radiación del reactor, no llega a ser radiactivo, es inerte químicamente y posee excelentes características para la transferencia de calor.

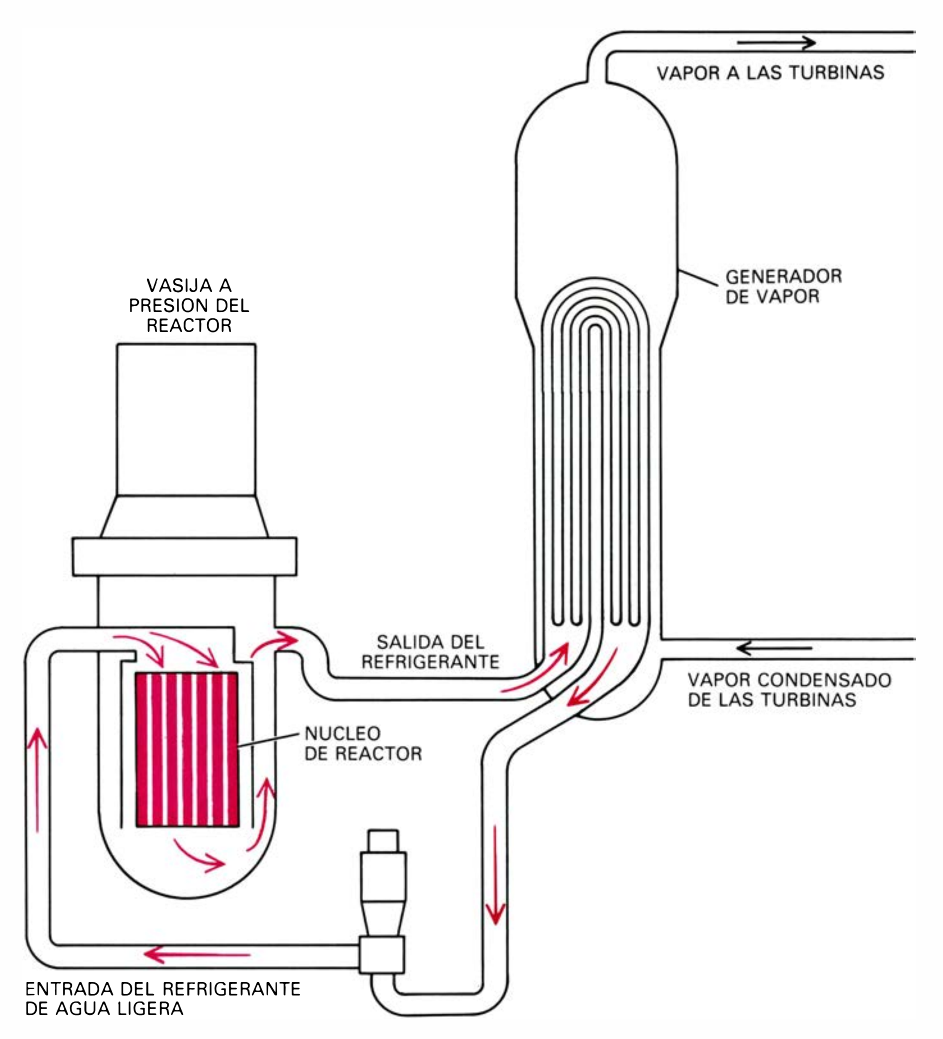
Las interesantes propiedades de los reactores HTGR las resumió Joseph M. Hendrie, director de la Comisión Reguladora Nuclear, en su informe ante el subcomité del Congreso en marzo de 1980. "Este tipo de reactores, expuso, alcanzan rendimientos tan buenos como las mejores centrales de combustible fósil y rendimientos más altos que los reactores refrigerados por agua. No solamente obtienen mayor rendimiento térmico, sino también una mejor utilización energética por cada kilogramo

de uranio extraído de la mina; en cifras, supera en un 15 o un 20 por ciento a los mejores cálculos realizados para los combustibles de los reactores más avanzados de agua ligera." Añadió que los HTGR "iban asimismo por delante en el ámbito de la seguridad. Son máquinas en las que no se necesita revolver demasiado ni muy deprisa cuando algo no acaba de funcionar, porque la estructura del núcleo es una gran masa de grafito, material muy estable a alta temperatura, de modo que, si se produce una caída de potencia o se para la circulación del refrigerante, hay tiempo suficiente para sentarse y pensar qué debe hacerse en cada caso".

La primera serie de reactores refrigerados por gas construidos en Inglaterra fueron los llamados reactores magnox; respondía el nombre al hecho de que las barras del combustible, que contenían uranio natural sin enriquecer, estaban envainadas en una aleación de magnesio. El núcleo del reactor, que albergaba muchas toneladas de grafito, se hallaba alojado a presión en una va-



NUCLEO DEL REACTOR Y GENERADORES DE VAPOR del HTGR; ambos están confinados en una enorme vasija de hormigón pretensado (PCVR). El PCVR de un reactor proyectado para generar 860 MWe mide 31 metros de diámetro por 29 de altura. (Un reactor de agua a presión de una capacidad ligeramente mayor se presenta en la misma escala en la ilustración de la página siguiente.) El núcleo del grafito del HTGR de 860 MWe ocupa un volumen cilíndrico de 8 metros de diámetro y 6,4 metros de altura. A través de unos 27.000 canales verticales del núcleo, circula gas helio a una presión de 74 kilogramos por centímetro cuadrado, impulsado por cuatro bombas principales, de las que sólo se ven dos en el corte transversal presentado en la figura. El helio sale del núcleo del reactor a una temperatura de 685 grados Celsius y entra por la base de los generadores de vapor, donde realiza un doble recorrido por un conjunto de tuberías rectas y helicoidales para la producción de vapor. El agua hierve en el interior de las tuberías cuando circula en dirección ascendente y se recalienta aún más al circular hacia abajo, para salir en forma de vapor recalentado a una temperatura de 540 grados Celsius y con una presión de 170 kilogramos por centímetro cuadrado. Aunque no aparecen en la figura, existen tres circuitos de refrigeración en los que intercambiadores de calor refrigerados por agua pueden extraer el calor del helio circulante cuando los generadores de vapor estén fuera de servicio. Al parar el reactor, los productos de fisión del núcleo desprenden calor a una tasa muy alta inicialmente, si bien decrece exponencialmente con el tiempo.



REACTOR DE AGUA LIGERA A PRESION, el tipo más común de reactores de energía nuclear, con 44 ya en servicio. Otros 24 reactores son del tipo de agua en ebullición, en los que el calor se elimina del núcleo por vapor de agua, y no por agua caliente a presión. En la figura se muestra el núcleo de un reactor de agua ligera a presión, de una potencia nominal de 1100 MWe. Está alojado en una vasija de acero a presión de unos 4,5 metros de diámetro, 12 metros de altura y de unos 15 a 28 centímetros de espesor. El proyecto prevé que la vasija funcione con una presión interna de 160 kilogramos por centímetro cuadrado. El agua de refrigeración sale del reactor a una temperatura de 321 grados Celsius y pasa a cuatro circuitos de generación de vapor, de los cuales sólo se indica uno en la figura. El vapor sale del generador a 282 grados Celsius y a una presión de 70 kilogramos por centímetro cuadrado. A esa temperatura y presión, el rendimiento térmico del sistema alcanza sólo del 32 al 33 por ciento, comparado con el 38,5 por ciento que se obtiene en un sistema tipo HTGR. (Ilustración realizada por George V. Kelvin.)

sija de acero muy cara, varias veces mayor que las vasijas necesarias para los reactores de agua ligera. Posteriormente, en 1958, los ingenieros franceses demostraron que podía reemplazarse la vasija de acero por una vasija de hormigón pretensado, cuyo tamaño permitiera alojar el sistema del reactor completo, incluidos los generadores de vapor. La vasija del reactor de hormigón pretensado, o PCRV (*P*restressed-*C*oncrete *R*eactor *V*essel), se mantiene en entera compresión por una red de cables de acero redundantes, sujetos a control y retensión, y hasta sustitución cuando es necesario. La hermeticidad contra las fugas se asegura con una capa de acero fijada a la pared interior del PCRV, que actúa sólo de membrana de cierre para el refrigerante. La capa de acero y las paredes del PCRV se refrigeran por agua que circula a través de tubos soldados a la superficie exterior de la vasija.

Los PCRV se adoptaron inmediatamente por todos los reactores refrigerados por gas franceses y británicos. El alto grado de seguridad ofrecido por la vasija de hormigón contribuyó a la decisión británica de construir una segunda generación de reactores, denominada reactores de gas avanzados (AGR, *A*dvanced *G*as *R*eactor), cerca de las ciudades. El combustible que se utilizó en esta segunda generación fue óxido de uranio, de un tipo cerámico, envainado en acero inoxidable, cambio que fue posible al emplear uranio ligeramente enriquecido. Con el nuevo combustible, los AGR pudieron operar a temperatura más alta que los reactores con combustible tipo magnox, y "quemar" más uranio 235 del combustible antes de que hubiera que reemplazar el combustible. Con temperaturas más altas, el rendimiento en generación de energía eléctrica se elevó de un 30 por ciento a un poco más del 40 por ciento.

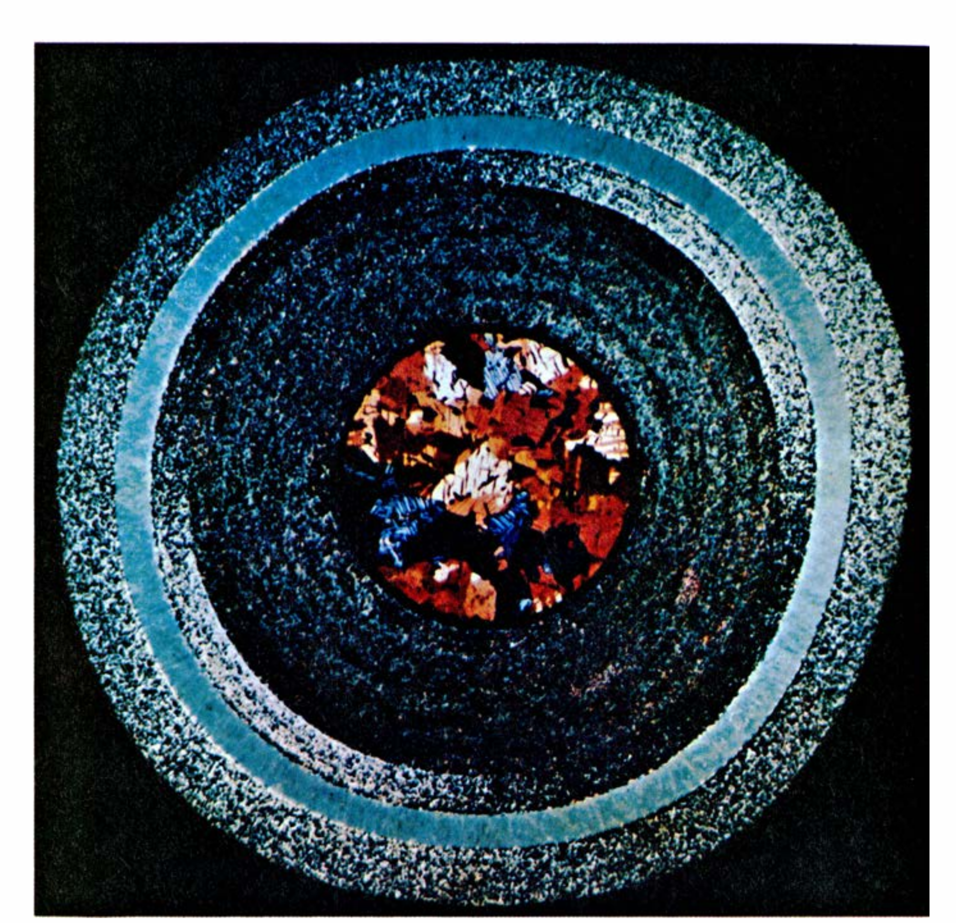
En Estados Unidos, la Comisión de Energía Atómica (oficina que antecedió al actual Departamento de Energía) demostró su interés hacia los reactores refrigerados por gas, en las décadas de 1950 y 1960, sufragando las investigaciones en varios aspectos avanzados relativos al reactor. Uno de los principales objetivos de la AEC era reducir la cantidad de uranio enriquecido por unidad de potencia eléctrica: en aquellas fechas no parecía que los suministros de uranio pudieran hacer frente a la demanda prevista. De ahí que la investigación se polarizó sobre aquellas cuestiones y puntos del reactor que tuvieran que ver con los reproductores o convertidores avanzados. Un reactor reproductor crea al menos un átomo de nuevo combustible por cada átomo de combustible consumido. Un reactor convertidor avanzado suele crear de 0,7 a un átomo de combustible por cada átomo consumido. Los reactores de agua ligera producen entre 0,5 y 0,6 átomos de combustible por cada átomo consumido. El reactor de alta temperatura refrigerado por gas y moderado por grafito puede clasificarse como un convertidor avanzado. Fue uno de los que sobrevivió tras el inevitable barrido de diseños. El HTGR goza del apoyo de la industria eléctrica, porque es competitivo en coste de capital con los reactores de agua ligera y porque utiliza el ciclo de combustible uranio-torio con un bajo consumo de uranio y, en consecuencia, bajos costos de combustible.

La continua evolución de la tecnología de los reactores de gas en Europa y en los Estados Unidos ha resultado en una mutua convergencia en dos importantes hechos, al menos, para la próxima etapa en el desarrollo de los reactores refrigerados por gas: se sustituye el dióxido de carbono por helio como refrigerante y se carga el núcleo del reactor con el combustible nuclear en un único sistema que evita la necesidad de una vaina de metal. Estos dos hechos se han comprobado no solamente en Peach Bottom y en Fort St. Vrain, sino también en dos reactores europeos. Los ingleses pusieron en funcionamiento un reactor de ensayos de 20 MW térmicos refrigerado por helio en el sur de Inglaterra desde 1965 a 1976. En Ale-

mania, un HTGR de 15 MWe (llamado AVR) ha venido generando energía eléctrica desde 1967, con una temperatura del gas a la salida muy alta, 950 grados Celsius. (La temperatura del agua a la salida del núcleo de un reactor de agua ligera a presión es de unos 320 grados Celsius.) Una central de 300 MWe, basada en la experiencia acumulada por el AVR, está hoy en fase de construcción en Alemania; su puesta en marcha se prevé para 1984 o 1985. En los Estados Unidos, el reactor de 330 MWe de Fort St. Vrain ha producido más de dos mil millones de kilowatt-hora desde 1978 y ha demostrado el aprovechamiento del combustible y las características de seguridad inherentes al diseño actual de los HTGR. Se ha sometido el reactor a ensayos con transitorios e incluso a la pérdida absoluta de la circulación forzada del refrigerante, sin que se apreciaran efectos adversos en el núcleo del reactor o en otros componentes principales del sistema.

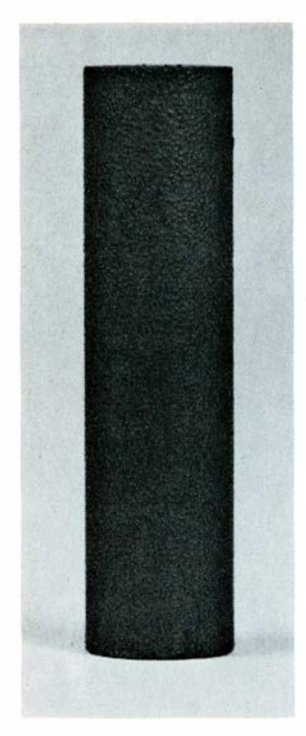
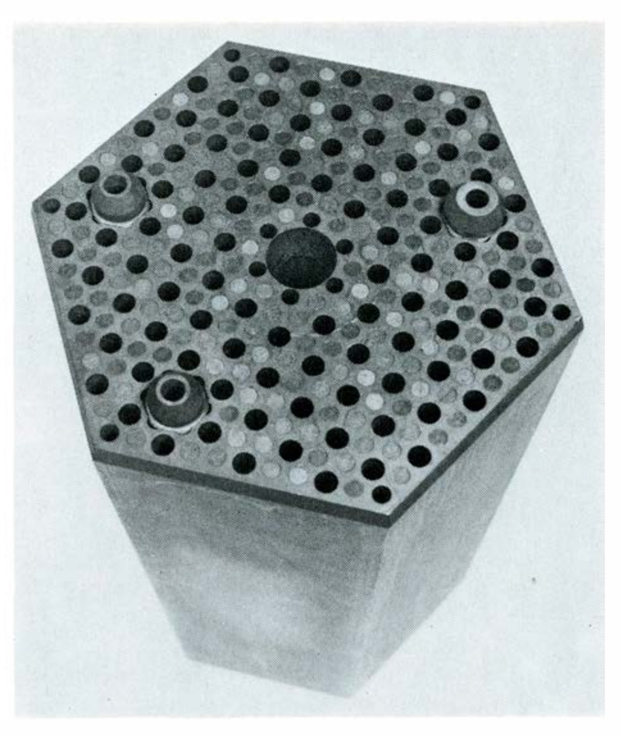
Basados en la experiencia del reactor de Fort St. Vrain, la General Atomic, en cooperación con la Asociación de Reactores Refrigerados por Gas (un consorcio de compañías eléctricas de los Estados Unidos) y del Departamento de Energía, ha desarrollado un diseño de referencia para un HTGR de 860 MWe. Se quiso un diseño simple y conservador, con especial énfasis en la seguridad y protección de la inversión de capital realizada. El núcleo del reactor está contenido en una vasija de hormigón pretensado con múltiples cavidades. El helio sale del núcleo a 685 grados Celsius (se rebajan los 812 grados Celsius del Fort St. Vrain) y pasa a través de cuatro circuitos principales de refrigeración, donde se genera vapor a una temperatura de 538 grados Celsius y a una presión de 170 kilogramos por centímetro cuadrado.

El helio atraviesa cada circuito de refrigeración, impulsado mediante un circulador movido por un motor eléctrico. (Los circuladores en Fort St. Vrain se mueven por vapor.) El núcleo está provisto también de un sistema auxiliar de refrigeración formado por tres circuitos, cada uno suficiente por sí mismo para proporcionar el 100 por cien de la refrigeración necesaria, cuando el helio esté a la presión normal de operación de 74 kilogramos por centímetro cuadrado, o el 50 por ciento de la refrigeración cuando la vasija esté despresurizada. El helio que pasa a través del sistema auxiliar de refrigeración se enfría con agua que circula impulsada por bombas eléctricas, que pueden





PARTICULA DE COMBUSTIBLE diseñada para el HTGR. Mide 0,75 milímetros de diámetro, el tamaño de un grano de arena. En la parte superior se presenta una sección transversal de la partícula, aumentada 150 veces. El combustible nuclear es el material cristalino del centro. Consiste en oxicarburo de uranio, que está enriquecido al 93 por ciento en su contenido del isótopo fisionable uranio 235. Sobre el núcleo se han depositado capas de carbón y carburo de silicio en un proceso a alta temperatura.

BARRA DE COMBUSTIBLE Y BLOQUE DE COMBUSTIBLE de un HTGR presentados a la derecha y a la izquierda de la figura, respectivamente. La barra de combustible, de unos 6 centímetros de longitud, consta de decenas de miles de partículas de combustible embebidas en una matriz de grafito. Cada bloque de combustible, de 35 centímetros de ancho y 80 centímetros de longitud, está formado por 1656 barras de combustible empaquetadas en un conjunto de forma hexagonal. Los numerosos canales vacíos del bloque se destinan al paso del flujo de helio. En el gran canal del centro, se aloja un mecanismo para insertar los bloques en el núcleo del reactor. El núcleo de un reactor de 860 MWe requiere 3512 bloques. Cada bloque, que pesa 122 kilogramos, contiene un promedio de 700 gramos de U 235 y unos 16 kilogramos de torio 232. En sus cuatro años de residencia en el reactor, cada uno de estos bloques producirá una energía equivalente a 2500 toneladas de carbón o 12.000 barriles de petróleo. Si los átomos de uranio 235 sin quemar y los uranio 233 creados a partir del torio se recuperasen y reciclaran, la energía equivalente del combustible original se elevaría a unas 11.000 toneladas de carbón o a 54.000 barriles de petróleo.

arrancar mediante generadores diesel, en caso necesario.

La combinación de un gas noble inerte como refrigerante y de un núcleo con estructura de grafito, muy resistente a la temperatura, permite generar vapor a las mismas altas temperaturas y presiones que en las plantas más modernas de producción de energía eléctrica a partir de combustibles fósiles. El rendimiento neto de generación de energía eléctrica del diseño de referencia del HTGR es del 38,5 por ciento, ligeramente por debajo del 39,2 por ciento obtenido en el reactor de Fort St. Vrain. Esta pequeña reducción se realizó para simplificar el sistema de generación de vapor y obtener más márgenes de operabilidad y seguridad.

Una propiedad fundamental del helio como refrigerante, un gas confinado en la vasija que no puede condensarse a fase líquida en el sistema, es que mantiene una relación lineal entre la temperatura y la presión; por ello, las lecturas instrumentales de presión y temperatura aportan comprobaciones independientes entre sí. Al no haber interfases líquido-gas, como ocurre en los reactores de agua hirviendo (y en los de agua a presión en ciertas condiciones de emergencia) una simple señal inequívoca –la presión– indica siempre la presencia y condición física del refrigerante. Puede tolerarse una rápida despresurización del sistema de refrigeración primario sin preocuparse por si se han formado vacíos que dejen parte del núcleo sin cubrir, como puede suceder cuando la presión la aporta agua por encima de su punto de ebullición atmosférico.

La experiencia del sistema de Fort St. Vrain ha comprobado varias ventajas importantes en la seguridad y en la operación de los HTGR. El personal de mantenimiento y operación ha recibido exposiciones a la radiación muy por debajo de los límites establecidos por la legislación para las centrales nucleares. No llegan a 10, de un total de varios cientos, los trabajadores que han recibido cantidades de radiación que hayan podido medirse.

El sistema de Fort St. Vrain ha respondido suavemente y sin alterarse a cambios inducidos en la carga por transitorios en el ciclo de generación de energía o por paradas temporales del equipo del interior de la central. El núcleo del HTGR tiene una gran capacidad de absorción de calor debido a su gran tamaño, a que desprende menos calor por unidad de volumen que los reactores de agua ligera y debido también a la masa enorme del núcleo, más de 1500 toneladas de grafito. Si el flujo del refrigerante se reduce o se interrumpe, el reactor responde lentamente a un fallo operacional inesperado, concediéndole al operador tiempo suficiente para realizar la acción correctora adecuada, horas en vez de segundos.

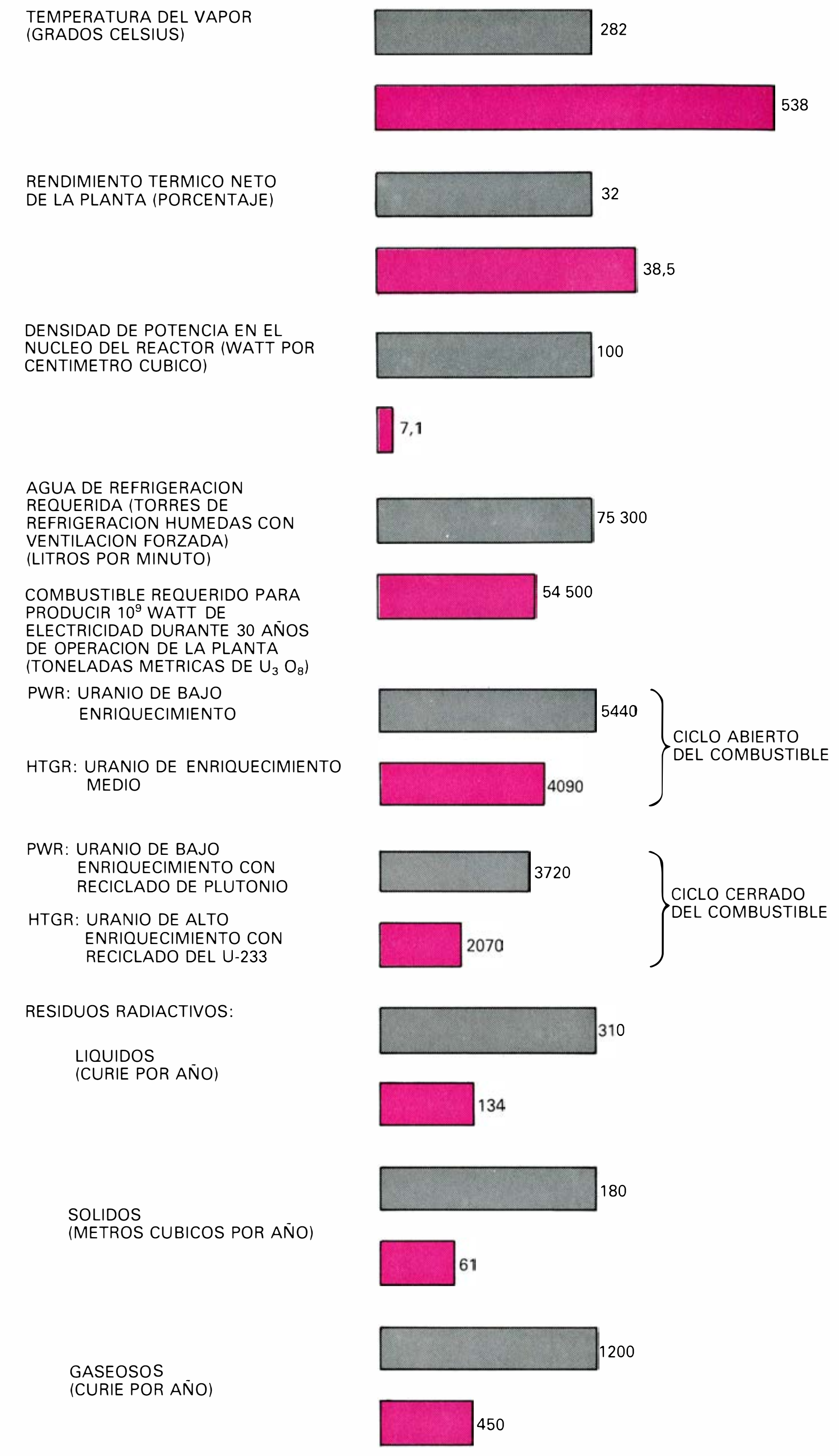
En Fort St. Vrain, cinco al menos de estos acontecimientos han interrumpido la circulación forzada de helio por extensos períodos, sin producir incrementos mensurables en la temperatura del núcleo o dañar de algún modo la planta o el combustible. El riesgo de producir daños en el reactor o en el núcleo por un error del operador está virtualmente eliminado. Gracias a la estabilidad térmica del HTGR, el sistema para detener la actividad del reactor, mediante la inserción de absorbentes de neutrones, y el sistema de refrigeración de emergencia pueden ser de un diseño muy simple. También hay suficiente margen de tiempo para hacer funcionar ambos sistemas manualmente, si las reglas y especificaciones nucleares lo permiten. A raíz del accidente de Three Mile Island, la Comisión Reguladora Nuclear exige ya la presencia permanente de un inspector experto, llamado el consejero técnico, en todas las plantas nucleares. El reactor de Fort St. Vrain es el único reactor exento de este mandato; no se requiere que el experto permanezca constantemente en la planta, sino tan sólo que pueda presentarse a la llamada en menos de una hora.

La vasija del reactor de hormigón pretensado constituye un rasgo importante desde el punto de vista de la seguridad nuclear. En primer lugar, la probabilidad de una rotura catastrófica del PCRV es tan remota, que los análisis de riesgos le niegan verosimilitud. Cada uno de los cables de acero, que dan al PCRV su fortaleza, son independientes y redundantes. La vasija se halla en un estado constante de compresión. En segundo lugar, el PCRV está diseñado para resistir una presión final de más del doble de la presión de operación, o sea, 170 kilogramos por centímetro cuadrado. Cualquier rotura de un cable de acero que pudiera ocurrir por una presión excesiva no haría más que producir una pequeña fuga de gas; tales fugas tienden a autocerrarse cuando se reduce ligeramente la presión. Tercero,

la despresurización total sólo puede ocurrir si falla una penetración de las tuberías, o cualquiera de las pequeñas tuberías de servicio que atraviesan las paredes del PCRV. Este hipotético fallo es un suceso de probabilidad extremadamente baja. Por si fuera poco, cada penetración de la vasija está equipada con limitadores de flujo que evitan una salida muy rápida del gas que pudiera causar daños estructurales en el núcleo o en el sistema de enfriamiento.

El HGTR también ofrece varias ventajas sobre la actual generación de reactores desde el punto de vista del medio ambiente. Al trabajar el HGTR con un rendimiento del 39 por ciento, frente al 33 por ciento de los reactores de agua ligera, desprende cerca de un 25 por ciento menos de calor residual a disipar en el medio ambiente que rodea la central. Si el calor se descarga en un río o lago cercano en forma de agua caliente, se reducen en la misma proporción los problemas resultantes de elevar la temperatura hasta límites peligrosos para la supervivencia del ecosistema acuático. Si se emplean torres de refrigeración para disipar el calor, éstas consumirán menos agua y pueden ser más pequeñas y más baratas. Si se emplea un embalse o estanque de refrigeración se puede situar una planta HTGR con un tercio más de megawatt que otra central de agua ligera sobre un embalse determinado sin exceder la temperatura especificada para el mismo. Si es necesario adoptar torres de refrigeración secas para cumplir con las regulaciones ambientales o para mantener inalteradas las fuentes de agua disponibles, la pérdida de capacidad de la planta en tiempo cálido será sólo la mitad con un HTGR que con cualquier central nuclear actual. En consecuencia, una planta HTGR puede instalarse en un lugar remoto, árido o semiárido, con un gravamen en su coste muy pequeño.

En todas las centrales se controla cuidadosamente el nivel de radiactividad de todas las descargas rutinarias. En virtud de su propio diseño, la planta HTGR tiene en los sistemas de flujo de fluidos del interior de la central menos radiactividad y a concentraciones más bajas que en un reactor de agua ligera. Además, los HGTR incorporan algunos dispositivos para asegurar que las descargas de radiactividad al medio ambiente sean nulas. Se espera que los procedimientos rutinarios de descontaminación produzcan pequeños volúmenes de residuos radiactivos líquidos de bajo nivel (menos de 7500 litros por



CARACTERISTICAS OPERATIVAS de un HTGR de 860 MWe (*en color*), comparadas con las de un reactor de agua a presión de la misma capacidad de generación de electricidad (*en gris*). El menor consumo de combustible del HTGR puede atribuirse, en parte, a su mayor rendimiento térmico y, en parte, al hecho de que, para cada átomo de U 235 consumido en el HTGR, se crean 0,7 átomos de combustible nuevo. El reactor de agua a presión no alcanza a producir 0,5 átomos de combustible nuevo por cada átomo que consume. Con un ciclo cerrado del combustible, ambos sistemas convierten cierta fracción de los átomos de U 238 o Th 232 en isótopos de plutonio o uranio, de los que algunos de ellos se consumen en el propio reactor antes de que se reemplace el combustible. Si el combustible agotado se reciclara (lo que iba en contra de la política de la administración Carter), sería preferible usar como combustible en un HTGR una mezcla de uranio altamente enriquecido (93 por ciento de U 235) y torio. Parte del torio se convertiría en U 233 fisionable, que sería recobrado y reciclado para reemplazar al U 235 en las posteriores cargas de combustible. Los menores volúmenes de residuos radiactivos de un HTGR resultan de su mayor rendimiento y de las ventajas refrigeradoras del helio sobre el agua.

año con una actividad total inferior a 150 curie). Este pequeño volumen puede enviarse fuera de la planta sin ningún problema, o retenerse dentro de la central. Los residuos radiactivos sólidos producidos por un HTGR deben ser inferiores a 56.000 litros por año. Cerca del 80 por ciento serán residuos de bajo nivel (papel, componentes de los filtros, resinas agotadas, etcétera) que están ligeramente contaminados; pueden enviarse al exterior de la planta en barriles para enterrarlos o quemarlos sin que incidan negativamente sobre el medio ambiente. El 20 por ciento restante serían residuos de nivel intermedio, consistentes fundamentalmente en bloques del reflector, que deberá reemplazarse periódicamente. Tales residuos pueden enviarse fuera de la planta para su almacenamiento definitivo a largo plazo en bidones blindados de 200 litros de capacidad.

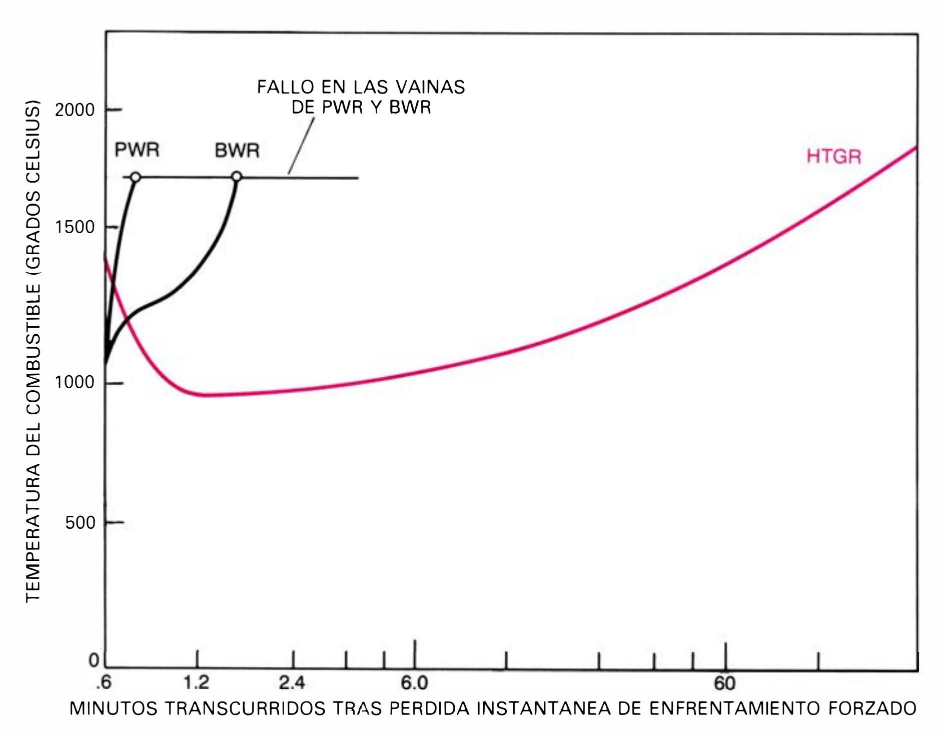
Los sistemas de purificación de helio y de recuperación de gases incorporados a las plantas standard HTGR deberán reducir los niveles radiactivos de los gases descargados por la planta varios órdenes de magnitud inferiores a las actuales especificaciones gubernamentales de cinco milirrem por año. El tritio (isótopo radiactivo del hidrógeno) que se genera en el primario del HTGR se elimina en el sistema de purificación de helio mediante un oxidante; éste convierte el tritio en agua tritiada, que posteriormente se solidifica y se manipula como un residuo sólido, fácilmente aislable mientras el tritio se desintegra. (El período de semidesintegración del tritio es de 12,26 años.)

El diseño del HTGR reúne ciertos rasgos característicos que simplifican las operaciones de mantenimiento y recarga del combustible. Por ejemplo, todo el refrigerante primario, helio, está confinado dentro de la vasija del reactor de hormigón pretensado, que proporciona por sí misma todo el blindaje necesario para protección del personal; así, la labor de mantenimiento puede realizarse en todo el edificio del reactor sin que tenga que pararse la central. Dado que el sistema de vapor del secundario está esencialmente libre de radiactividad, todos los equipos componentes del ciclo de vapor fuera del PCRV, incluidos los equipos de la planta de turbinas, pueden operarse y mantenerse como los de una central típica de combustibles fósiles. Por otro lado, y puesto que el flujo de vapor en las turbinas de un HTGR es sólo el 60 por ciento del flujo en las turbinas de una planta de la misma potencia de un reactor de agua ligera, todos los equipos relacionados con el ciclo de vapor y de agua de alimentación son más pequeños y, por ende, más fáciles de mantener. En general, el mantenimiento, reparación y manipulación cuesta menos en una planta HTGR que en otra de agua ligera, porque el helio, al contrario que el agua, es inerte, no corrosivo y no se convierte en radiactivo.

Otra gran ventaja del empleo de un gas como refrigerante es su transparencia, que hace posible inspeccionar visualmente muchas zonas dentro del PCVR. El blindaje a la radiación inherente en el diseño del PCRV posibilita muchas tareas de inspección y mantenimiento mientras el reactor está en funcionamiento, lo que reduce el tiempo de parada del reactor a ese objeto.

Casi todos los componentes estructurales del PCRV, como los cables verticales y la envoltura de cables circulares de la vasija, pueden inspeccionarse visualmente mientras el reactor está en funcionamiento. Algunos cables seleccionados se hallan bajo control ininterrumpido para detectar cualquier cambio de tensión que revelaría un deterioro de su función. En caso necesario, se reemplaza cualquier parte de la estructura. Todas las superficies exteriores del hormigón, menos las contiguas a los portones de las barras de control, pueden también inspeccionarse visualmente sin tener que parar la central. Los portones de las barras de control y las superficies que rodean las penetraciones de los mecanismos de las barras de control pueden observarse fácilmente durante la recarga del combustible.

Experimentos recientes de recarga del combustible en Fort St. Vrain han demostrado la fácil manipulación de los elementos de combustible tipo bloque del HTGR. Cerca de 240, un sexto, de los elementos combustibles se retiraron del núcleo y se reemplazaron por elementos combustibles nuevos. Los restantes 1240 elementos no se movieron del lugar que ocupaban. El equipo que realizó la recarga recibió niveles tan bajos de radiación que, para medirlos, fue necesario emplear un monitor de microrrem. Por extrapolación de los datos existentes, se pudo calcular que la exposición total integrada en hombre-rem para toda la operación de la



SEGURIDAD INHERENTE A UN HTGR mostrada en la figura, donde se comparan las temperaturas en el núcleo de un HTGR, en el núcleo de un reactor de agua a presión (PWR) y en el núcleo de un reactor de agua en ebullición (BWR), tras un hipotético accidente de pérdida de refrigerante o de pérdida de la circulación forzada. En los reactores refrigerados por agua, la reacción nuclear se para automáticamente por la pérdida del agua que actúa como moderador. En el HTGR la reacción nuclear debe pararse por la inserción de las barras de control que absorben neutrones. En el momento de detención, la desintegración de los productos de fisión en el combustible desprende calor a una velocidad equivalente al 7 por ciento de la potencia térmica neta del reactor. El calor desprendido desciende al uno por ciento en dos horas, y al 0,5 por ciento a las 24 horas. En los reactores refrigerados por agua, en ausencia de refrigeración de emergencia, la temperatura de la vaina de los elementos combustibles subiría en menos de dos minutos a 1650 grados Celsius, estropeando las vainas. En el HTGR la masa de moderador de grafito absorbería el calor desprendido por los productos de fisión, de modo que los 1650 grados Celsius no se alcanzarían hasta después de una hora. Una temperatura suficientemente elevada para producir daños en el núcleo de grafito (alrededor de 2200 grados Celsius) precisaría diez horas o más sin refrigeración forzada.

recarga, realizada con el reactor funcionando a plena potencia, no llegó a los cinco hombre-rem. Las regulaciones federales actuales limitan la dosis a cinco rem por individuo y año.

Cada elemento combustible del HTGR es un bloque de grafito con una sección transversal hexagonal de 35 centímetros de ancho por 80 centímetros de longitud. El bloque está perforado longitudinalmente con 72 canales de refrigeración y 138 canales ciegos para alojar el combustible. El grafito constituye una opción ideal como moderador y como material estructural, porque su fortaleza aumenta al hacerlo la temperatura. En el diseño de referencia, los bloques de grafito de elementos combustibles se agrupan en columnas de ocho. Esta disposición segmentada axialmente facilita la fabricación, la manipulación y la recarga.

La adopción de esta configuración en bloques, tan favorable, ha sido posible por el desarrollo de una partícula de combustible especialmente recubierta. El núcleo central de cada partícula es una microesfera de oxicarburo de uranio (convenientemente enriquecido en uranio 235) de 0,25 milímetros de diámetro. En torno a cada esfera se colocan capas muy finas de carbón, carbón pirolítico y carburo de silicio, aplicadas a alta temperatura que originan una partícula muy encapsulada con un diámetro total de 0,75 milímetros. Para las partículas de torio se utiliza una técnica de encapsulamiento análoga. Esta técnica asegura la contención de los productos de fisión. La integridad estructural de estas pequeñas esferas se comprueba, en lotes de 2000 unidades, exponiéndolas a un flujo de radiación que simula las condiciones internas del reactor. El proceso de producción de estas partículas, semiautomático, va acompañado de un procedimiento riguroso de ensayos para alcanzar un severo control de calidad.

Aunque la existencia de condiciones graves e imprevisibles en una región del reactor pudiera causar el fallo de la envoltura de las partículas de combustible y se produjera la liberación de productos de fisión, el fallo quedaría limitado a la zona donde esto ocurriera. En la mayoría de los reactores donde la vaina con los elementos combustibles está situada a lo largo de todo el núcleo del reactor, un fallo operacional que cause la rotura de una pequeña sección de la vaina libera todos los productos de fisión existentes a lo largo de toda la barra del combustible. El comportamiento de los elementos combustibles en Fort St. Vrain ha satisfecho cumplidamente las expectativas del diseño. Incluso se puede decir que la liberación de productos de fisión ha sido muy inferior a los valores previstos. En resumen, las barreras de contención opuestas a los productos de fisión de los elementos combustibles del HTGR han demostrado poseer un grado elevado de fiabilidad.

AÑO	DISPONIBILIDAD DEL SISTEMA NUCLEAR DE GENERACIONDE VAPOR (PORCENTAJE)	DISPONIBILIDAD DE LA PLANTA (PORCENTAJE)	FACTOR TOTAL DE CAPACIDAD DE LA PLANTA (PORCENTAJE)
1967	81	78	69
1968	88	88	82
1969	86	84	67
1970	95	95	88
1971	90	87	78
1972	71	71	58
1973	95	94	78
1974	96	95	70
NUCLEO 1 PROMEDIO	85	83	73
NUCLEO 2 PROMEDIO	89	88	74
PROMEDIO DURANTE LA VIDA DE LA PLANTA	88	86	74

FIABILIDAD DE LA PRIMERA PLANTA HTGR, proyectada por la General Atomic, la Estación Generadora de energía No. 1 de Peach Bottom. Hasta qué punto es fiable esa central nos lo indica la estadística de la tabla. Aparte de las paradas programadas y del tiempo perdido por causas ajenas al funcionamiento del propio reactor, el HTGR estuvo disponible para suministrar vapor para la generación de energía el 88 por ciento del tiempo. Al alcanzar el 75 por ciento de su capacidad nominal de generación de electricidad durante sus siete años y medio de actividad, el reactor de Peach Bottom sobrepasó la cifra típica del 66 por ciento alcanzado por los reactores de agua ligera que administran las compañías eléctricas.

Las propiedades que caracterizan a los HTGR permiten utilizar una amplia variedad de ciclos de combustible nuclear. El ciclo que se ha estudiado y ensayado con mayor intensidad es el ciclo uranio-torio en el que se utiliza uranio totalmente enriquecido (93 por ciento de U 235) como material fisionable primario y torio (Th 232) como material "fértil". En el reactor, el torio absorbe neutrones y se convierte finalmente en el isótopo fisionable uranio 233, que puede reciclarse para sucesivas recargas de combustible. El reactor de Fort St. Vrain está cargado con combustible de uranio enriquecido al 93,5 por ciento en U 235, en combinación con torio. Merced al diseño de la planta, se puede utilizar uranio totalmente enriquecido o con un enriquecimiento medio (alrededor del 20 por ciento de U 235). Los costes del ciclo del combustible en el HTGR, con las restricciones actuales en el reciclado y reproceso del combustible, pueden parangonarse con los de otras centrales comerciales. Mientras no se cambie de política, el combustible agotado se almacena de forma permanente sin proceder a la recuperación, ni del U 235 aún sin utilizar, ni del U 233 o del plutonio, creados a lo largo del funcionamiento del reactor. Este ciclo del combustible se denomina habitualmente ciclo abierto.

Si un HTGR opera en ciclo abierto con combustible de uranio-torio, con uranio totalmente enriquecido, consumirá, durante su período de vida útil cifrado en 40 años, cerca del 20 por ciento menos de uranio que un reactor de agua ligera equivalente. Si los dos tipos de reactores operaran en ciclo cerrado con una recuperación total del uranio y del plutonio, el HTGR consumiría el 50 por ciento menos de uranio. Consecuentemente, el HTGR ofrece la ventaja de ahorrar cantidades importantes de uranio, tanto en ciclo cerrado como en ciclo abierto del combustible, con tal de que el reactor esté diseñado para operar con combustible totalmente enriquecido, La importancia de los ahorros potenciales de uranio se pone de manifiesto si atendemos al coste total del combustible durante toda la vida útil de una central nuclear de producción de energía eléctrica, que viene a ser igual al coste inicial de la planta.

En los últimos seis años se han cancelado las órdenes de construcción de 55 plantas nucleares ya proyectadas. Desde hace solamente seis años, las compañías de producción de energía eléctrica de los Estados Unidos han mostrado interés en construir 10 centrales HTGR, al menos. Cuando el reactor de Fort St. Vrain funcione a plena potencia, y haya demostrado la excepcional seguridad y fiabilidad que le pronostican los ingenieros, las compañías eléctricas contemplarán favorablemente las plantas tipo HTGR, una vez estén listas para cursar nuevas órdenes de compra de centrales nucleares eléctricas.

La desintegración del protón

Se sabe que el protón tiene una vida media de al menos 10^{20} veces la edad del universo, pero la teoría indica que quizá no viva eternamente. Si no es inmortal, toda la materia ordinaria acabará por desintegrarse

Steven Weinberg

El descubrimiento de la radiactividad por Antoine Henri Becquerel en 1896 rebatió la creencia de que todos los átomos eran permanentes e inmutables. Posteriormente se supo que las partículas energéticas que habían sido detectadas por Becquerel se emitían cuando los núcleos de los átomos de una sustancia radiactiva se desintegraban espontáneamente en otros núcleos atómicos. Aunque interesante, esta inestabilidad nuclear parecía una rareza, una propiedad de sólo ciertos elementos pesados como el uranio y el radio. Se creía que los núcleos de elementos como el hidrógeno y el oxígeno eran absolutamente estables.

Hay ya razones teóricas para sospechar que todos los núcleos atómicos acaban desintegrándose y que, por tanto, toda la materia es, en pequeña medida, radiactiva. En una desintegración de esta clase, uno de los dos tipos de partículas del núcleo atómico, un protón o un neutrón, se transformaría espontáneamente en partículas energéticas muy diferentes de las que constituyen los átomos ordinarios. Hasta el núcleo más ligero, el de hidrógeno, que consta de un simple protón, estaría sujeto a esta desintegración.

Existen muchas indicaciones, empezando por la enorme edad de la Tierra, que señalan que la materia no puede ser altamente disipante. Si la materia ordinaria se desintegra, debe hacerlo muy despacio, tan lentamente que se necesitarán experimentos de una escala extraordinariamente grande para detectar la desintegración. Becquerel descubrió la desintegración radiactiva de los núcleos de uranio en un cristal de sales de uranio que pesaba tal vez algunos gramos; para observar la mucho más débil radiactividad asociada a la desintegración del protón, será preciso manejar muchas toneladas de material. A pesar de ello, hay en marcha varias investigaciones experimentales centradas en la desintegración del protón.

Para entender qué es lo que se dilucida en estos experimentos, resulta útil preguntarse primero si existen razones para que algo del mundo sea eterno. Por ejemplo, se cree todavía que el electrón es absolutamente estable. ¿Qué principios físicos le impiden desintegrarse en otras partículas? Entendiendo la estabilidad de partículas como el electrón, podremos juzgar si hay principios que impidan la desintegración de núcleos atómicos ordinarios.

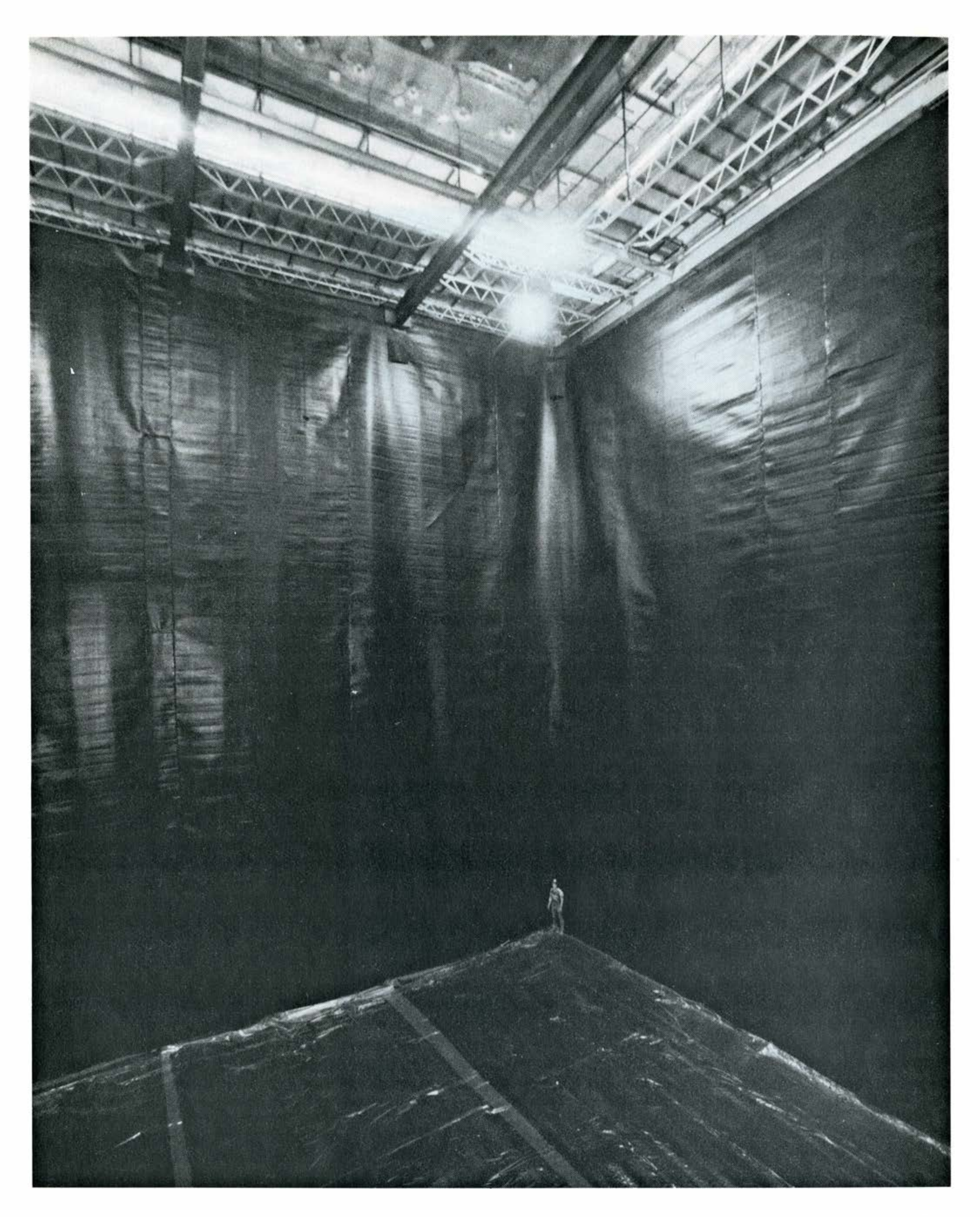
La experiencia en física de partículas elementales enseña que cualquier proceso de desintegración que pueda imaginarse ocurre espontáneamente, a menos que esté prohibido por alguna de las leyes de conservación de la física. Establece una de ellas que el valor total de cierta cantidad, como la energía o la carga eléctrica, nunca puede cambiar. Aun cuando un proceso de desintegración no se produzca directamente por una de las interacciones fundamentales de las partículas elementales, y no esté prohibido por una ley de conservación, se producirá a través de alguna sucesión más o menos complicada de emisiones y absorciones de partículas. Así, al considerar si una partícula es o no estable debemos preguntarnos si su desintegración violaría alguna ley de conservación.

La aplicación de la ley de conservación de la energía no entraña mayor dificultad. Requiere simplemente que la masa de la partícula que se desintegra (o la energía equivalente a esta masa) sea mayor que la masa total de los productos de desintegración. (No es suficiente que ambas sean iguales ya que cierta masa debe convertirse en la energía cinética de los productos de la desintegración.) Por tanto, un buen punto de partida para juzgar la estabilidad de cualquier partícula es hacer una lista de todas las partículas de menos masa en las que, en principio, podría desintegrarse.

Consideremos el electrón. Que se sepa, sólo hay unas pocas clases de partículas cuya masa sea inferior a la del electrón. La más familiar de ellas es el fotón, el cuanto de luz, cuya masa se cree que es exactamente cero. Hay fuertes razones teóricas para abogar también por un cuanto de radiación gravitacional, el gravitón, de masa también nula. Finalmente, existen varias especies de partículas llamadas neutrinos, que en algunos aspectos son similares al electrón; se emiten en un tipo familiar de radiactividad conocida como desintegración beta, el tipo de radiactividad descubierto por Becquerel en 1896. En general, se ha supuesto que los neutrinos tienen masa cero, pero la determinación de su masa es actualmente objeto de un intenso esfuerzo teórico y experimental. De todas formas, no hay ninguna duda de que al menos un tipo de neutrino tiene una masa inferior a aproximadamente una diezmilésima parte de la masa del electrón.

¿Por qué no se desintegra pues el electrón en neutrinos y fotones, por ejemplo? La respuesta es que, aunque tal desintegración satisfaría la ley de conservación de la energía, violaría otra ley de conservación, la de la carga eléctrica. Benjamin Franklin fue el primero en observar que la cantidad neta de carga eléctrica (positiva menos negativa) nunca aumenta ni disminuye, aunque cargas de signo opuesto pueden separarse o recombinarse. Los electrones tienen una carga eléctrica negativa definida, mientras que todas las partículas más ligeras en las que el electrón podría desintegrarse (el fotón, el gravitón y los neutrinos) resultan tener carga eléctrica nula. La desintegración de un electrón significaría la destrucción de una cantidad negativa definida de carga eléctrica y está, por tanto, estrictamente prohibida.

Veamos ahora de qué forma se aplicarían estas leyes de conservación a la



GRAN CAVIDAD a unos 600 metros bajo tierra en la mina de sal de Morton, al este de Cleveland, que en los próximos meses se llenará de 10.000 toneladas de agua, dispuesta para la búsqueda de la desintegración de un protón o un neutrón ligado en un núcleo atómico. La cavidad mide unos $18 \times 24 \times 21$ metros. La desintegración de uno cualquiera de los $2,5 \times 10^{33}$ protones y neutrones de la región central de agua originará partículas de alta velocidad. Cuando una partícula cargada se mueve dentro de un medio transparente con una velocidad mayor que la de la luz en dicho medio, emite un cono de luz azul llamado radiación de Cerenkov. Esta radiación es un estruendo óptico análogo al estruendo sonoro originado por un avión que vuela más velozmente que el sonido en el aire. La luz de Cerenkov la detectarán 2400 tubos fotomultiplicadores, que se instalarán por las paredes de la cavidad. El experimento se realiza bajo tierra para reducir el número de partículas de alta energía procedentes del espacio que podrían confundirse con productores de desintegración de un protón o de un neutrón ligado. En el estudio participa la Universidad de California en Irvine, la de Michigan y el Brookhaven.

desintegración de las dos clases de partícula que constituyen el núcleo atómico. Por el momento consideremos solamente la más ligera de ellas, el protón, y dejemos el neutrón para más adelante. El protón tiene carga eléctrica positiva, igual en magnitud y opuesta en signo a la del electrón; tampoco puede desintegrarse, por tanto, en neutrinos, fotones o gravitones. El protón, sin embargo, es unas 1820 veces más pesado que el electrón, y existen varias partículas más ligeras que tienen también carga positiva. El protón podría desintegrarse en estas otras partículas sin violar la conservación de la energía ni la de la carga eléctrica. Por ejemplo, el electrón tiene una antipartícula llamada positrón, con la misma masa que el electrón pero con una carga eléctrica positiva igual a la del protón. (Para cada tipo de partícula existe una antipartícula con la misma masa, pero con valores opuestos de otras propiedades, como la carga eléctrica. A propósito, el positrón es estable por la misma razón que lo es el electrón.) No hay nada en las leyes de conservación de la energía o de la carga que prohíba al protón desintegrarse en un positrón y un número cualquiera de fotones y neutrinos.

Otro candidato a producto de desintegración del protón es el antimuón. El muon es una partícula similar en muchos aspectos al electrón; dotado de la misma carga, pesa, empero, 210 veces más. (El muon se desintegra en un electrón y neutrinos.) El antimuón tiene la misma carga que el protón y sólo alrededor de una novena parte de su masa. Un protón podría desintegrarse, por tanto, en un antimuón más otras partículas neutras ligeras, como el fotón y los neutrinos.

Existe aún otro posible producto de la desintegración del protón, un mesón, es decir, un miembro del grupo de partículas inestables de masa intermedia entre la del electrón y la del protón. Las leyes de conservación de la energía y de la carga permitirían que el protón se desintegrara en un mesón cargado positivamente y un neutrino o bien en un mesón neutro y un positrón, por ejemplo. Cualquiera de estos procesos de desintegración llevaría a la total destrucción del átomo de hidrógeno. En un elemento más pesado cambiaría su naturaleza química y liberaría energía en cantidades mucho mayores que las liberadas en la radiactividad ordinaria.

¿Por qué no se observa en ninguna parte que la materia se esté desintegrando como resultado de tales procesos? Según parece, este problema lo planteó, por primera vez, Hermann Weyl en 1929. Los positrones, muones y mesones eran desconocidos por entonces y, en consecuencia, no podían sospecharse los hipotéticos esquemas de desintegración del protón más arriba esbozados. A pesar de ello, Weyl se sintió intrigado por la estabilidad de la materia; probablemente se preguntó por qué en un átomo los protones no absorben a los electrones orbitales originando, por ejemplo, la desintegración del átomo de hidrógeno en una lluvia de fotones. Weyl sugirió que podría explicarse la estabilidad de la materia si

LOSA DE HORMIGON en la mina de Soudan en Minnesota sensibilizada para detectar sucesos que indiquen la desintegración del protón. Unos 3450 detectores de partículas han sido empotrados en la losa, una mole de 31 toneladas. Los contadores detectarán directamente los productos de la desintegración del protón. El experimento lo dirigen investigadores de la Universidad de Minnesota y del Argonne.

hubiese dos tipos de carga eléctrica, una acarreada por el electrón y otra por el protón. Si cada tipo de carga se conservara por separado, estaría prohibida la aniquilación mutua de un protón con un electrón. La propuesta de Weyl no despertó entonces mucha atención.

La cuestión fue abordada de nuevo por E. C. G. Stueckelberg en 1938 y por Eugene P. Wigner (en una nota a pie de página) en 1949. Propusieron lo que se ha convertido en el punto de vista convencional, esto es: además de la energía y la carga eléctrica, existe otra propiedad conservada de la materia que, desde entonces, se ha llamado número bariónico. Los bariones (del griego *barys,* pesado) constituyen una familia de partículas que incluye al protón y muchas partículas más pesadas que el protón, tales como el neutrón y las partículas altamente inestables llamadas hiperones. A todos los bariones se le asigna un número bariónico +1, y todas las partículas más ligeras, entre las que se cuenta el fotón, el electrón, el positrón, el gravitón, el neutrino, el muon y los mesones, tienen un número bariónico cero. Para un átomo o cualquier otro sistema compuesto de partículas, el número bariónico es la suma de los números bariónicos de las partículas constituyentes. De ello se sigue que cualquier colección de partículas más ligeras que el protón tiene un número bariónico cero. La ley de conservación del número bariónico es la aserción de que el número bariónico total no puede cambiar. La desintegración del protón en una colección de partículas más ligeras significaría la conversión de un estado cuyo número bariónico es +1 en un estado cuyo número bariónico es cero y, por tanto, la desintegración está prohibida.

Una antipartícula tiene un número bariónico opuesto al de la partícula correspondiente. El antiprotón, por ejemplo, tiene un número bariónico −1; es un antibarión. Un protón y un antiprotón pueden aniquilarse mutuamente sin violar la conservación del número bariónico; el protón y el antiprotón tiene un número bariónico total +1 más −1, o sea cero, y pueden, por consiguiente, convertirse en una lluvia de mesones o fotones. Así pues, la conservación del número bariónico no exige que cada protón sea inmortal, sino más bien que los protones no se desintegren espontáneamente en la materia ordinaria, donde no hay antiprotones.

Hasta ahora me he limitado a la desintegración del protón. Mas el núcleo de la mayoría de los átomos no consta sólo de protones, sino también de neutrones. ¿Qué podemos decir acerca de la posible desintegración del neutrón? El neutrón es un barión con carga eléctrica cero y con una masa ligeramente mayor que la del protón. Para ser más precisos, la masa del neutrón es un poco mayor que la masa del protón más la del electrón. Esta relación sugiere un posible modo de desintegración para el neutrón: podría producir un protón, un electrón y algunas partículas neutras de masa nula. Evidentemente, la energía puede conservarse en este proceso. Lo mismo ocurre con la carga eléctrica, ya que las cargas del protón y del electrón se cancelan mutuamente. El número bariónico también se conserva, pues el neutrón y el protón tienen cada uno número bariónico +1, mientras que las demás partículas poseen número bariónico cero.

Un neutrón libre (no ligado en un núcleo atómico) se desintegra exactamente de esta forma: produce un protón, un electrón y un antineutrino. El semiperíodo de un neutrón libre, que es el tiempo necesario para que la mitad de los neutrones de una gran muestra se desintegren, es de unos 10 minutos. Los neutrones de ciertos núcleos atómicos, como el núcleo de tritio (el isótopo pesado del hidrógeno constituido por un protón y dos neutrones), pueden desintegrarse también en protones; este es el proceso de desintegración beta. En la mayoría de los núcleos, sin embargo, los neutrones no se desintegran debido a que se necesitaría demasiada energía para crear un protón en medio de las fuerzas electrostáticas repulsivas generadas por los demás protones del núcleo. En tales núcleos, los neutrones son tan estables como los protones.

Cabría la posibilidad de que un neutrón ligado en un núcleo pudiera desintegrarse de alguna otra forma en la que no se conservara el número bariónico. Por ejemplo, podría producir un positrón y un mesón cargado negativamente o bien un electrón y un mesón positivamente cargado. El descubrimiento de tal desintegración en un núcleo, que de otro modo se mostraría estable, sería tan significativo como el descubrimiento de la desintegración del protón. En efecto, los experimentos para poner de manifiesto la desintegración del protón buscan también la desintegración de neutrones ligados. Sin embargo, dado que la desintegración de neutrones libres es ya bien conocida, las pruebas experimentales sobre la conservación del número bariónico son conocidas como experimentos de desintegración del protón.

En los últimos años se ha convertido en un aserto comúnmente aceptado el que bariones y mesones están formados por partículas más fundamentales llamadas quarks. Un barión está constituido por tres quarks, un antibarión por tres antiquarks y un mesón por un quark y un antiquark. El electrón, el muon y los neutrinos pertenecen a la familia de partículas llamadas leptones, que no están formadas por quarks y, en efecto, no muestran ningún indicio de tener estructura interna. Sobre esta base, el número bariónico de un sistema de partículas es justamente un tercio del número neto de quarks, esto es, un tercio de la diferencia entre el número de quarks y el de antiquarks. La conservación del número neto de quarks es equivalente a la conservación del número bariónico.

Tal vez el lector escéptico se sienta algo insatisfecho por recurrir a la conservación del número bariónico para explicar la estabilidad del protón y del neutrón ligado. Esta sensación estaría, en mi opinión, justificada. El número bariónico surgió como instrumento de contaje para justificar la no observación de las desintegraciones protónicas ni de otras relacionadas; no tiene ninguna otra significación conocida. En este aspecto, el número bariónico difiere de la carga eléctrica, que posee un significado dinámico directo: una carga eléctrica crea campo eléctrico y magnético y, a su vez, estos campos actúan sobre la carga. La teoría de la electricidad y el magnetismo no tendría sentido alguno si la carga eléctrica no se conservara; en cambio, no se conoce ningún argumento dinámico análogo para la conservación del número bariónico.

En efecto, hay pruebas empíricas en contra de la existencia de algún tipo de campo (llamémosle campo bariotrópico) que esté relacionado con el número bariónico, como sí lo está el campo electromagnético con la carga eléctrica. La Tierra contiene del orden de 4×10^{51} protones y neutrones y posee, por tanto, un número bariónico enorme. Si la Tierra fuese una fuente de campo bariotrópico debería esperarse que dicho campo atrajera o repeliera a los protones y neutrones de los cuerpos ordinarios sobre la superficie terrestre. Una fuerza bariotrópica podría distinguirse de la fuerza gravitacional, ya que la fuerza gravitacional que la Tierra ejerce sobre un cuerpo es proporcional a la masa de dicho cuerpo, mientras que la fuerza bariotrópica sería proporcional a su número bariónico. Cuerpos de igual masa compuestos de distintos elementos pueden tener números ba-

PARTICULA	MASA (MeV)	CARGA ELECTRICA	NUMERO BARIONICO	PRINCIPAL MODO DE DESINTEGRACION
FOTON (γ)	0	0	0	NINGUNO CONOCIDO
NEUTRINO (ν)	0?	0	0	NINGUNO CONOCIDO
ANTINEUTRINO ($\bar{\nu}$)	0?	0	0	NINGUNO CONOCIDO
ELECTRON (e^-)	.511	−1	0	NINGUNO CONOCIDO
POSITRON (e^+)	.511	+1	0	NINGUNO CONOCIDO
MUON (μ^-)	105.7	−1	0	$\mu^- \rightarrow e^- + \nu + \bar{\nu}$ MASA: 105.7 → .511 + 0 + 0 CARGA: −1 → −1 + 0 + 0 NUMERO BARIONICO: 0 → 0 + 0 + 0
ANTIMUON (μ^+)	105.7	+1	0	$\mu^+ \rightarrow e^+ + \nu + \bar{\nu}$ MASA: 105.7 → .511 + 0 + 0 CARGA: +1 → +1 + 0 + 0 NUMERO BARIONICO: 0 → 0 + 0 + 0
MESONES PI (π^+)	139.6	+1	0	$\pi^+ \rightarrow \mu^+ + \nu$ MASA: 139.6 → 105.7 + 0 CARGA: +1 → +1 + 0 NUMERO BARIONICO: 0 → 0 + 0
(π^0)	135	0	0	$\pi^0 \rightarrow \gamma + \gamma$ MASA: 135 → 0 + 0 CARGA: 0 → 0 + 0 NUMERO BARIONICO: 0 → 0 + 0
(π^-)	139.6	−1	0	$\pi^- \rightarrow \mu^- + \bar{\nu}$ MASA: 139.6 → 105.7 + 0 CARGA: −1 → −1 + 0 NUMERO BARIONICO: 0 → 0 + 0
MESONES *K* (K^+)	493.7	+1	0	$K^+ \rightarrow \mu^+ + \nu$ MASA: 493.7 → 105.7 + 0 CARGA: +1 → + 1+ 0 NUMERO BARIONICO: 0 → 0 + 0
(K^0_S)	497.7	0	0	$K^0_S \rightarrow \pi^+ + \pi^-$ MASA: 497.7 → 139.6 + 139.6 CARGA: 0 → +1 + −1 NUMERO BARIONICO: 0 → 0 + 0
(K^0_L)	497.7	0	0	$K^0_L \rightarrow \pi^0 + \pi^0 + \pi^0$ MASA: 497.7 → 135 + 135 + 135 CARGA: 0 → 0 + 0+ 0 NUMERO BARIONICO: 0 → 0 + 0 + 0
(K^-)	493.7	−1	0	$K^- \rightarrow \mu^- + \bar{\nu}$ MASA: 493.7 → 105.7+ 0 CARGA: −1 → −1 + 0 NUMERO BARIONICO: 0 → 0 + 0
PROTON (p)	938.3	+1	+1	NINGUNO CONOCIDO
ANTIPROTON ($\bar{p}$)	938.3	−1	−1	NINGUNO CONOCIDO
NEUTRON (n)	939.6	0	+1	$n \rightarrow p + e^- + \bar{\nu}$ MASA: 939.6 → 938.3 + .511 + 0 CARGA: 0 → +1 + −1 + 0 NUMERO BARIONICO: +1 → +1 + 0 + 0
ANTINEUTRON ($\bar{n}$)	939.6	0	−1	$\bar{n} \rightarrow \bar{p} + e^+ + \nu$ MASA: 939.6 → 938.3 + .511 + 0 CARGA: 0 → −1 + +1 + 0 NUMERO BARIONICO: −1 → −1 + 0+ 0
HIPERON Λ (Λ^0)	1115.6	0	+1	$\Lambda^0 \rightarrow p + \pi^-$ MASA: 1115.6 → 938.3 + 139.6 CARGA: 0 → +1 + −1 NUMERO BARIONICO: +1 → +1 + 0
ANTIHIPERON Λ ($\bar{\Lambda}^0$)	1115.6	0	−1	$\bar{\Lambda}^0 \rightarrow \bar{p} + \pi^+$ MASA: 1115.6 → 938.3 + 139.6 CARGA: 0 → −1 + +1 NUMERO BARIONICO: −1 → −1 + 0

riónicos que difieran en casi un 1 por ciento. Varios experimentos altamente precisos (empezando por los de Roland von Eötvös en 1889) demuestran que la atracción de los cuerpos por la Tierra es, de hecho, estrechamente proporcional a su masa, no a su número bariónico. En 1955, T. D. Lee, de la Universidad de Columbia, y C. N. Yang, del Instituto de Estudios Avanzados de Princeton, demostraron a partir de un análisis de estos experimentos que cualquier fuerza bariotrópica entre dos partículas nucleares tendría que ser mucho más débil que la fuerza gravitacional, la cual es a su vez casi 40 órdenes de magnitud más débil que la fuerza electromagnética. No puede descartarse de forma absoluta que el número bariónico desempeñe un papel dinámico como el de la carga eléctrica, pero el argumento de Lee y Yang hace que tal papel aparezca como muy improbable.

La conclusión de que el número bariónico no tiene un papel dinámico no implica que el número bariónico no se conserve. En efecto, desde mediados de los años 30 los físicos se han familiarizado con otras cantidades que tampoco parecen disfrutar de un significado dinámico y, no obstante, se conservan, al menos en ciertos contextos. Entre estas cantidades están las denominadas extrañeza, isospín y conjugación de carga. Por ejemplo, a los protones y neutrones se les asigna extrañeza cero, a algunos hiperones se les asigna extrañeza −1 y a ciertos mesones, los llamados mesones *K*, se les asigna extrañeza +1. La conservación de la extrañeza fue introducida como una regla de con-

PRINCIPALES MODOS de desintegración de las partículas, gobernados por las leyes físicas de conservación fundamentales. La ley de conservación de la energía exige que la masa de la partícula que se desintegra sea tan grande, al menos, como la masa total de los productos de la desintegración. La ley de conservación de la carga eléctrica exige que la carga de la partícula que se desintegra sea igual a la carga total de los productos de la desintegración. La ley de conservación del número bariónico establece que el número bariónico de la partícula que se desintegra debe ser igual a la suma de los números bariónicos de los productos de la desintegración. El protón y muchas partículas más pesadas que él tienen un número bariónico de +1. Todas las partículas más ligeras que el protón tienen un número bariónico cero. La lista de partículas está dada en orden creciente de masa. La masa viene dada en función de su energía equivalente, en unidades de millones de electronvolt (MeV). La carga se expresa tomando como unidad la carga del protón. Bajo cada modo de desintegración, hay un contaje de las tres cantidades conservadas. La desintegración de un protón en una colección de partículas más ligeras significaría la transformación de un estado cuyo número bariónico es +1 en otro estado cuyo número bariónico es cero y, por tanto, la desintegración está prohibida por la conservación del número bariónico. Si se descubre que el protón se desintegra, la conservación del número bariónico dejará de ser universalmente cierta.

taje para explicar la observación de que un mesón K o un hiperón no pueden producirse aisladamente en colisiones de núcleos atómicos ordinarios, pero sí pueden producirse asociados, ya que un mesón K y un hiperón tienen extrañeza total cero. Después de la introducción de la idea del número bariónico, pareció plausible durante años que la conservación del número bariónico era otra de esas reglas de contaje no dinámicas que resulta ser universalmente satisfecha.

Este punto de vista sobre las leyes de conservación ha cambiado radicalmente con el desarrollo de las modernas teorías sobre las interacciones de las partículas elementales. Las teorías en cuestión describen todas las fuerzas conocidas entre las partículas elementales (aparte de la gravitación) de forma muy similar a como se exponía el electromagnetismo en la vieja teoría de las interacciones puramente electromagnéticas; esta última, llamada electrodinámica cuántica, fue desarrollada durante las décadas de 1930 y 1940. Se sostiene hoy que existen 12 campos similares al campo electromagnético de la electrodinámica cuántica. Estos campos son los ocho campos gluónicos que proporcionan las fuerzas nucleares fuertes que mantienen unidos a los quarks dentro de los bariones y los mesones, y los cuatro campos electrodébiles que, de una forma unificada, proporcionan las fuerzas nucleares débiles responsables de la desintegración beta y, a la vez, las del propio electromagnetismo. Existen 12 leyes de conservación asociadas, similares a la de la conservación de la carga eléctrica, para las cantidades denominadas color, isospín electrodébil e hipercarga electrodébil. (El color es una propiedad de los quarks que no tiene nada que ver con el color visual; la carga eléctrica es una combinación particularmente ponderada de la hipercarga y el isospín electrodébiles.) Al contrario que el número bariónico, estas cantidades conservadas detentan un significado físico directo: son las partículas que acarrean estas cantidades las que dan origen a los campos gluónicos y electrodébiles y éstos, a su vez, ejercen fuerza sobre tales partículas. La fuerza depende de los valores de las 12 cantidades acarreadas por la partícula.

A la vez que han aparecido nuevas leyes de conservación, las viejas leyes de conservación no dinámicas han quedado, en cierto sentido, desbancadas. Por ejemplo, la teoría moderna de las interacciones nucleares fuertes está tan estrechamente constreñida por la conservación del color (y por otros principios), que no hay forma de incluir los tipos de complicaciones que se necesitarían para violar la conservación de la extrañeza. Se podría intentar introducir interacciones fundamentales que no conservaran la extrañeza, pero siempre resulta posible redefinir qué es lo que se entiende por extrañeza de suerte que se siga conservando. Así, la conservación de la extrañeza no se concibe ya como un principio fundamental, cual la conservación de la energía o de la carga, sino como una consecuencia de la teoría detallada de las interacciones fuertes y, en particular, de la ley genuinamente fundamental de la conservación del color. Dado que la conservación de la extrañeza no es un principio fundamental de la física, no existe una razón general por la que deba respetarse fuera del dominio de las interacciones fuertes. En efecto, desde que se descubrió la extrañeza se sabe que ésta no se conserva en las interacciones nucleares débiles.

Las demás leyes de conservación no dinámicas han sido desbancadas de forma similar; ya no se consideran como leyes fundamentales de conservación, al mismo nivel que la conservación de la energía o la carga, sino como meras consecuencias matemáticas de la estructura de las teorías actuales de las interacciones de las partículas elementales. Una lista de las leyes de conservación consideradas hoy fundamentales incluiría la conservación de las 12 cantidades asociadas a las fuerzas electrodébiles y fuertes, la conservación de cantidades como la energía y el momento, que están asociadas de forma similar con las fuerzas gravitatorias, y la conservación del número bariónico, el cual no está asociado, que se sepa, a ninguna fuerza.

Este hecho por sí sólo debería hacernos desconfiar acerca de la conservación del número bariónico: el número bariónico no necesita ser conservado como la energía, la carga, el color y otras cantidades similares para tener teorías razonables sobre las interacciones de las partículas elementales. Además, existen indicaciones positivas de que la conservación del número bariónico no es exacta. Una de estas indicaciones es la proporcionada por la teoría moderna de las interacciones electrodébiles. Gerard 't Hooft, de la Universidad de Utrecht, ha demostrado que en esta teoría hay ciertos efectos sutiles, no representables por un número finito de emisiones y absorciones de partículas elementales, que inducen, aunque en proporciones extraordinariamente bajas, procesos en los que no se conserva el número bariónico. Estos procesos son demasiado lentos para ser detectados, pero importa notar que surgen precisamente debido a que la conservación del número bariónico no está relacionada con ningún tipo de campo bariotrópico; ningún efecto similar podría producir la no conservación de cantidades que, como la carga eléctrica, están relacionadas con campos de fuerza.

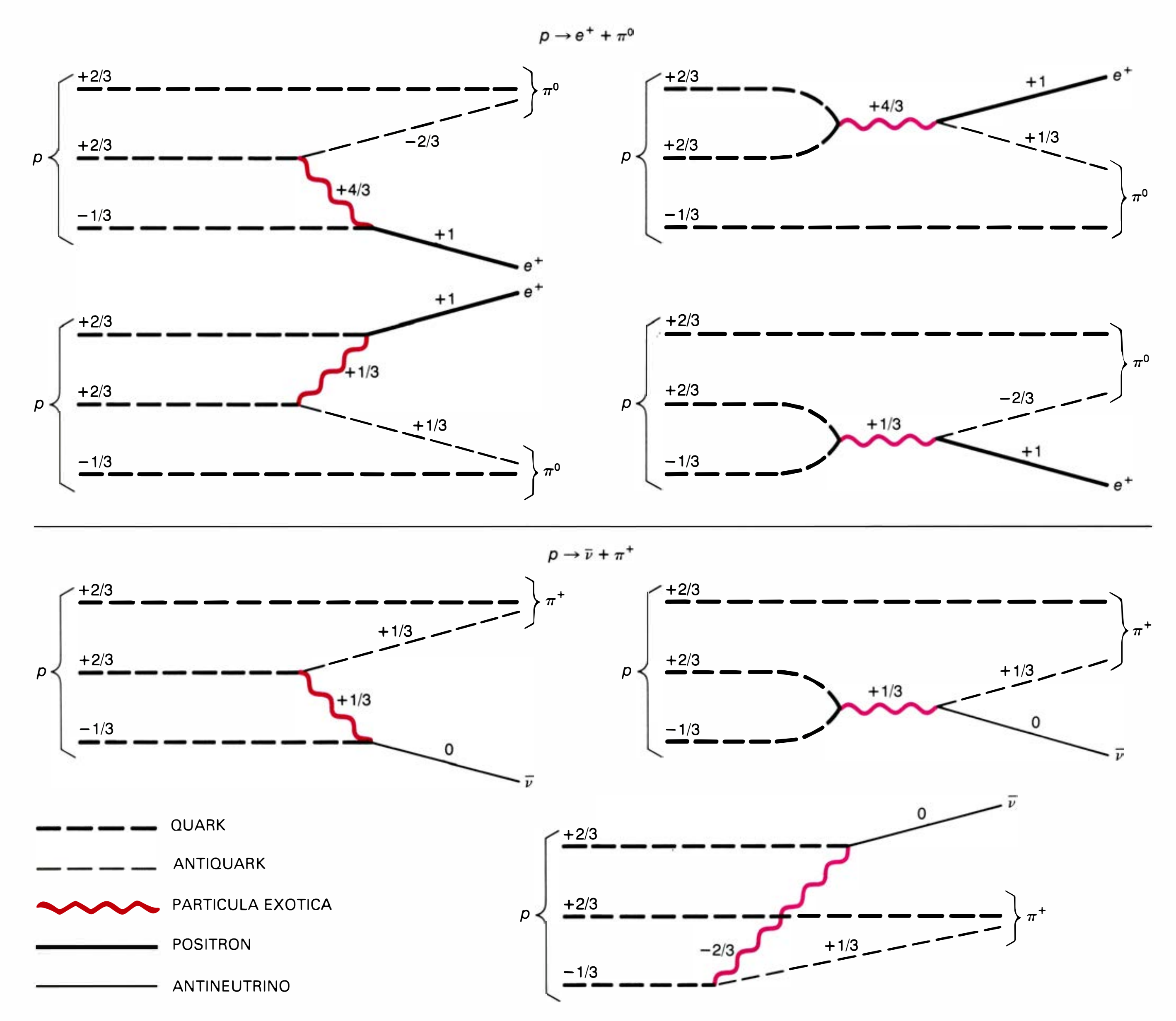
Otro indicio de que el número bariónico quizá no se conserve nos viene de la cosmología. Podemos suponer, aunque sólo sea por razones estéticas, que el universo empezó con cantidades iguales de materia y antimateria y, en consecuencia, con igual número de bariones y antibariones. En esta hipótesis, el universo habría empezado con un número bariónico total igual a cero. Si el número bariónico se conservara, debería seguir siendo cero. Casi todos los protones y neutrones se habrían aniquilado en colisiones con antiprotones y antineutrones y el universo contendría hoy sólo un tenue caldo de fotones y neutrinos, sin estrellas ni planetas ni científicos.

Existe la posibilidad de que el universo empezara con un exceso de materia sobre antimateria, de tal forma que algo quedaría después de la aniquilación entre partículas y antipartículas. También es posible (aunque suele considerarse improbable) que, de alguna manera, materia y antimateria se hayan separado y que nosotros estemos viviendo en una zona de número bariónico positivo en un universo con número bariónico total cero. Sin embargo, si el número bariónico no se conserva existe una posibilidad más atractiva, a saber, que el universo empezó con igual cantidad de materia y antimateria y que el exceso actual de materia con número bariónico positivo se debe a procesos físicos que han violado la conservación del número bariónico. (Desde el experimento de James H. Christenson, James W. Cronin, Val L. Fitch y René Turlay, de la Universidad de Princeton, en 1964, se sabe que no existe una simetría exacta materia-antimateria, la cual requiriría que los procesos en los que se crean antibariones ocurrieran en la misma proporción que aquellos en los que se crean bariones.) Estas consideraciones, junto con los argumentos de Lee y Yang en contra de la existencia de fuerzas bariotrópicas, llevaron a algunos físicos teóricos (entre los que cabe citar al físico ruso Andrei D. Sakharov y a mí mismo) a sugerir en la década de 1960 que el número bariónico podría no conservarse exactamente.

Consideraciones cosmológicas motivaron también al menos uno de los experimentos sobre la desintegración del protón realizados en esta época, el de T. Alväger, I. Martinson y H. Ryde, de la Universidad de Estocolmo y del Instituto Nobel. En años recientes muchos físicos teóricos han logrado construir esquemas sobre la producción de bariones en el universo primitivo.

Cualquier sugerencia sobre la posible no conservación del número bariónico ha de confrontarse inmediatamente con el hecho de que la materia ordinaria es muy estable. Maurice Goldhaber, del Laboratorio Nacional de Brookhaven, ha hecho notar que "sabemos en nuestros huesos" que la vida media del protón es superior a los 10^{16} años. Si la vida media fuese más corta, los 10^{28} protones del cuerpo humano estarían desintegrándose a un promedio de más de 10^{12} protones por año, es decir, 30.000 desintegraciones por segundo, y seríamos nuestra propia amenaza de enfermedad.

Por supuesto, se puede establecer un límite más restrictivo a la vida media del protón investigando activamente su desintegración. El primer experimento de este tipo fue realizado en 1954 por Frederick Reines y Clyde L. Cowan, Jr., en Los Alamos, y por Goldhaber. Usaron alrededor de 300 litros de hidrocarburo de centelleo, un material en el que las partículas cargadas energéticas, producidas en la desintegración de un protón, generarían un destello de luz detectable. Como habría de ocurrir con todos los experimentos de desintegración del protón que le siguieron, el aparato se instaló bajo tierra para resguardarlo de los rayos cósmicos. (Las partículas energéticas de estos rayos pueden causar sucesos que podrían confundirse con la desintegración de un protón.) Con esta precaución observaron sólo algunos destellos por segundo, casi todos los cuales podían ser atribuidos a rayos cósmicos que penetraron profundamente bajo tierra. Reines, Cowan y Goldhaber concluyeron que la vida media del protón debía ser superior a alrededor de 10^{22} años.



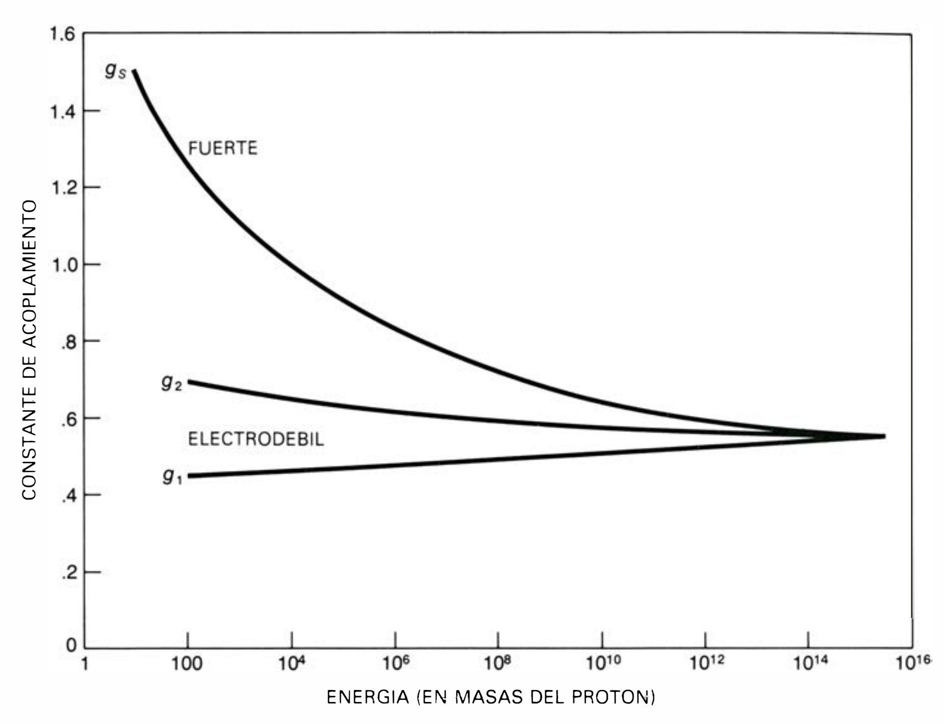
DESINTEGRACION DEL PROTON por emisión y absorción de una partícula exótica pesada. Lo más probable es que dé origen a un positrón (e^+) y un mesón neutro (π^0) o a un antineutrino ($\bar{\nu}$) y un mesón positivo (π^+). Se cree que el protón es una partícula compuesta formada por tres constituyentes llamados quarks; un mesón consta de un quark y un antiquark. La emisión o absorción de una partícula exótica produce la transformación de dos quarks en un antiquark y un positrón (*las cuatro reacciones superiores*) o en un antiquark y un antineutrino (*las tres reacciones inferiores*). Los quarks, positrones, antineutrinos y partículas exóticas están indicados por su carga eléctrica. La carga se conserva en las siete reacciones, que en esencia representan todas las maneras en las que la emisión y absorción de una sola partícula exótica podría provocar la desintegración de un protón en un mesón y un positrón o en un mesón y un antineutrino. Varios tipos de partículas exóticas pueden dar lugar a la desintegración del protón. Estas se distinguen por su carga eléctrica (más o menos 1/3, más o menos 2/3 y más o menos 4/3) y por su momento angular intrínseco o spin. El spin puede ser igual a cero o a 1.

Experimentos posteriores realizados por varios físicos han incrementado gradualmente la cota inferior empírica sobre la vida media del protón. La investigación más elaborada hasta este momento, cuyos resultados hayan sido publicados, fue realizada por un consorcio de investigadores de la Universidad de Case Western Reserve, la Universidad de Witwatersrand y la Universidad de California en Irvine. Usaron 20 toneladas de hidrocarburo de centelleo a una profundidad de unos 5 kilómetros en una mina de oro sudafricana, desde 1964 hasta 1971. Un análisis reciente de los datos ha dado como resultado que la vida media del protón o del neutrón ligado rebasa los 10^{30} años.

Se trata de una vida media en verdad muy dilatada. Para hacer una comparación, la edad actual del universo se estima en tan sólo 10^{10} años. Puede esperarse observar la desintegración de partículas con vidas medias tan largas debido a que los procesos de desintegración radiactiva operan estadísticamente: una muestra de partículas con una vida media de t años no sobrevive durante t años, desintegrándose luego todas al unísono; por el contrario, una fracción $1/t$ del número total de partículas se desintegrarán durante el primer año, $1/t$ de las partículas restantes se desintegrará durante el año siguiente, y así sucesivamente. La cota inferior para la vida media del protón no se establece observando un protón durante mucho tiempo y esperando que se desintegre, sino observando los 10^{31} protones y neutrones que hay en las 20 toneladas de material de centelleo a lo largo de varios años y esperando que se desintegren algunas docenas.

Es la larga vida media del protón la que nos llevó a la idea de la conservación del número bariónico. ¿Cómo podría el protón vivir tanto, de no haber una ley de conservación que lo hiciera persistir eternamente? En los últimos años se ha esbozado una respuesta.

Recordemos que la teoría moderna de las interacciones débiles, electromagnéticas y fuertes está altamente restringida, tanto que, por ejemplo, es imposible que las interacciones fuertes violen la conservación de la cantidad denominada extrañeza. Y si exceptuamos el pequeño efecto de 't Hooft, tampoco puede complicarse hasta el punto de permitir que la conservación del número bariónico se viole, a menos que se introduzcan nuevas partículas con exóticos valores de carga, color y demás propiedades. Estas partículas tendrían que ser cualitativamente distintas de las actualmente conocidas.



INTENSIDAD de las tres fuerzas fundamentales (fuerte, electromagnética y débil) que gobiernan las interacciones de las partículas elementales. Esta intensidad varía con la energía de los procesos en los que las fuerzas son medidas. La fuerza electromagnética actúa entre partículas con carga eléctrica, la interacción fuerte mantiene unidos a los quarks para formar el protón y otras partículas. La fuerza débil es responsable de ciertas desintegraciones radiactivas. La fuerza electromagnética y la débil se han unido en una teoría llamada teoría electrodébil, que propone dos tipos de fuerza electrodébil. La intensidad intrínseca de la fuerza fuerte y de las fuerzas electrodébiles vienen dadas por tres constantes de acoplamiento adimensionales (designadas g_s para la fuerza fuerte y g_1 y g_2 para las fuerzas electrodébiles). Aunque se les llama constantes, varían lentamente con la energía. Si las fuerzas tienen un origen común, existirá algún valor de la energía para el que las tres constantes de acoplamiento posean un valor común. Bajo hipótesis bastante generales, la energía de unificación se ha calculado en 10^{15} veces la masa del protón.

Si admitimos partículas exóticas del tipo adecuado, la desintegración del protón es posible. Las leyes usuales de conservación de la carga, color, etcétera, indican que lo que se necesita es una partícula con una carga eléctrica de $+4/3$, $+1/3$ o $-2/3$ veces la carga del protón; debe tener también un momento angular intrínseco (spin) igual a 0 o 1 y color idéntico al del antiquark. Tal partícula podría producirse, por ejemplo, al convertirse un quark en un antiquark y, destruirse, al convertirse otro quark en un antileptón (un positrón, un antimuón o un antineutrino); de esta forma, los tres quarks que constituyen el protón podrían dar lugar a un antileptón y a un mesón formado por el quark sobrante y el antiquark.

Estas partículas exóticas habrían de ser muy pesadas; de lo contrario, deberían haberse detectado ya. Si fueran suficientemente pesadas, sólo podrían ser emitidas y reabsorbidas con dificultad induciendo la desintegración del protón a un ritmo muy bajo. Cabe, pues, explicar la larga vida de un protón sin recurrir a una ley de conservación fundamental que asegure que el protón vive eternamente, lo que, por otro lado, ha abierto la posibilidad de que no viva eternamente.

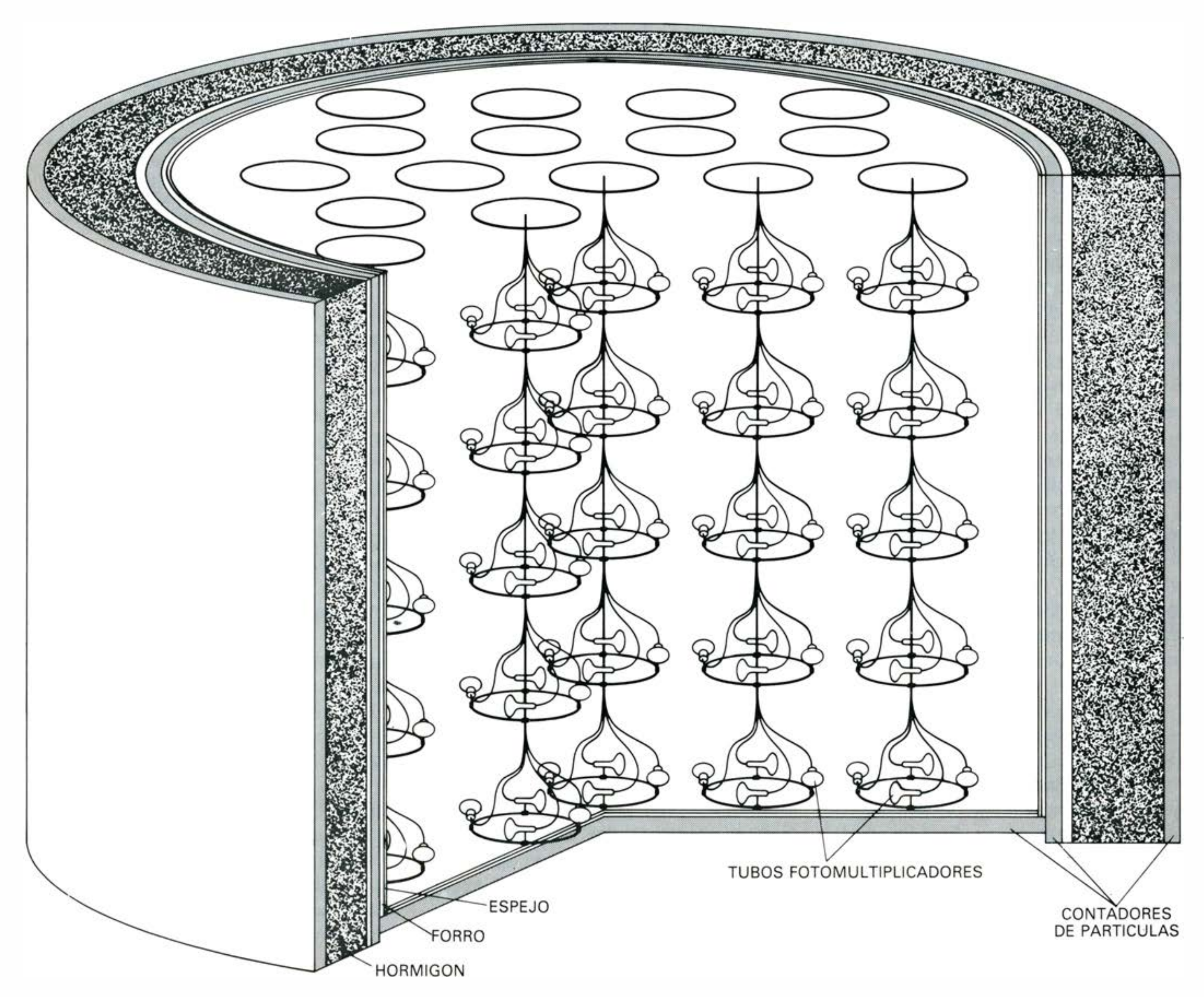
¿Hasta dónde habría de llegar la "pesantez" de las partículas exóticas para justificar la larga vida media del protón? Suponiendo que las partículas exóticas interactúen más o menos como los fotones, podemos estimar groseramente que una vida media del protón superior a 10^{30} años requiere una partícula exótica con una masa mayor que unas 10^{14} veces la del protón. Esta masa, de impresionante magnitud, es mayor que la que nadie pueda esperar producir con los aceleradores hoy por hoy imaginables. A pesar de ello, hay al menos dos razones para sospechar que estas partículas tan pesadas puedan de hecho existir.

La primera razón tiene que ver con el fenómeno de la gravitación, que hasta ahora hemos dejado de lado en nuestras consideraciones. La teoría de la relatividad general de Einstein constituye una teoría satisfactoria de las interacciones gravitacionales entre las partículas a todas las energías accesibles experimentalmente. Sin embargo, debido a fluctuaciones cuánticas, la teoría no es válida a energías muy altas, del orden de 10^{19} veces la masa del protón. Esta masa se conoce como la masa de Planck, en honor de Max Planck, quien en 1900 notó que una masa de este or-

den aparecería de forma natural en cualquier intento de combinar su teoría cuántica con la teoría de la gravitación. La masa de Planck es, hablando de un modo muy simplista, la energía a la que la fuerza gravitacional entre partículas se hace mayor que las fuerzas electrodébiles y fuertes. Para evitar una incoherencia entre la mecánica cuántica y la relatividad general, nuevos ingredientes deben entrar en la física a energías del orden (o por debajo) de 10^{19} veces la masa del protón.

La otra razón para esperar nuevos grados de libertad a tan altas energías está relacionada con las interacciones electrodébiles y fuertes. La teoría moderna de estas fuerzas contiene tres parámetros, denominados constantes de acoplamiento. Una de las constantes de acoplamiento, llamada g_s, describe la fuerza con que los campos gluónicos de las fuerzas fuertes interactúan con las partículas que llevan las cantidades conservadas que hemos llamado color; las otras dos constantes de acoplamiento, llamadas g_1 y g_2, describen la fuerza con que los campos electrodébiles interactúan con las partículas que llevan las correspondientes cantidades, hipercarga electrodébil e isospín electrodébil. Nos gustaría creer que todas estas interacciones tienen algún origen común, en cuyo caso todas las constantes de acoplamiento deberían ser del mismo orden de magnitud. Pero ello está en aparente contradicción con el hecho evidente de que las interacciones fuertes son fuertes; el valor que las medidas dan para g_s es mucho mayor que el que dan para g_1 o g_2.

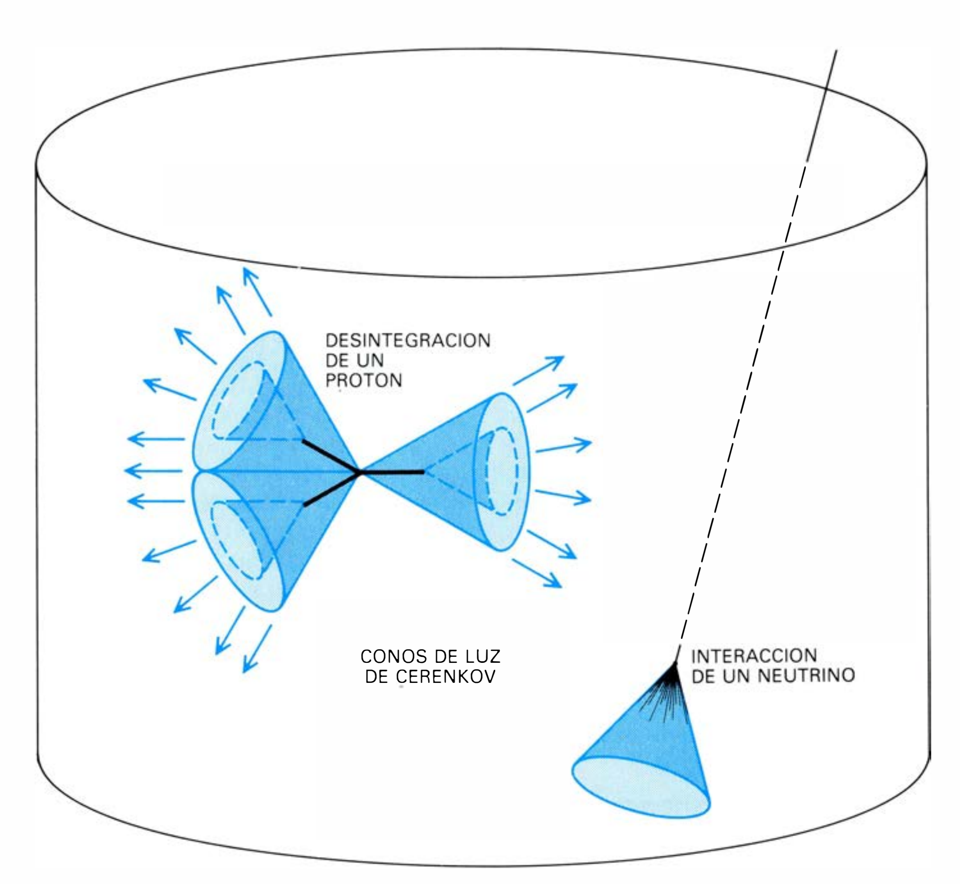
Una solución a esta dificultad fue propuesta en 1974 en la Universidad de Harvard por Howard Georgi, Helen R. Quinn y yo mismo. Era sabido desde el trabajo de 1954 de Murray Gell-Mann, del Instituto de Tecnología de California, y Francis E. Low, del Instituto de Tecnología de Massachusetts, que las constantes de acoplamiento dependían algo de la energía de los procesos físicos en las que se medían. En 1973, cálculos realizados independientemente por H. David Politzer, de Harvard, y David Gross y Frank Wilczek, de Princeton, demostraron que la constante de acoplamiento fuerte g_s disminuía lentamente al aumentar la energía. La mayor de las dos constantes electrodébiles, g_2, también decrece, pero más lentamente, mientras que la menor de ellas, g_1, aumenta con la energía. Lo que Georgi, Quinn y yo propusimos fue que la escala de energías a la que las interacciones fuertes se unificaban con las electrodébiles era enormemente alta, tan alta que la lentísima disminución de la constante de acopla-



DETECTOR DE MIL TONELADAS en fase de construcción en la mina Silver King, cerca de Park City, Utah. El detector tendrá 800 tubos fotomultiplicadores distribuidos por todo el volumen de agua. Los tubos fotomultiplicadores estarán sujetos por aros de alambre y la cámara se rodeará de contadores de partículas adicionales y de un grueso recubrimiento de hormigón. El volumen de agua sujeto a control contendrá unos 6×10^{32} protones y neutrones. El detector lo están construyendo investigadores que proceden de la Universidad de Harvard, la Universidad de Purdue y la Universidad de Wisconsin.

miento fuerte al aumentar la energía y la aún más lenta variación de las constantes de acoplamiento electrodébiles haría converger todas ellas hacia un valor esencialmente común a esta escala superalta de energías. Más concretamente, bajo hipótesis bastante generales (en esencia, que las interacciones fuertes y electrodébiles se unifican por un conjunto de simetrías del tipo conocido matemáticamente como grupo "simple", que no existen estadios intermedios de unificación y que las partículas elementales de spin 1/2 forman esquemas más o menos similares a los esquemas usuales de los leptones y los quarks) encontramos que la escala de energías a la que las interacciones fuertes y electrodébiles se unifican es del orden de 10^{15} a 10^{16} veces la masa del protón.

Cualquier teoría que unifique las interacciones fuertes y las electrodébiles, tratando de la misma forma a quarks y leptones, debe incluir nuevos tipos de partículas para completar el esquema y, como ya he argumentado anteriormente, no hay razón para pensar que las interacciones de estas nuevas partículas conserven el número bariónico. (La escala de energías de 10^{15} a 10^{16} veces la masa del protón que Georgi, Quinn y yo calculamos es lo suficientemente alta para que las interacciones que no conservan el número bariónico producidas por partículas exóticas de esta masa no pronostiquen una vida media del protón que contradiga el actual límite inferior experimental de 10^{30} años. Nosotros estimamos, muy grosso modo, una vida media de 10^{32} años.)



LA RADIACION DE CERENKOV no es sólo una señal de una posible desintegración de un protón, sino también de partículas de alta energía procedentes del espacio. En efecto, la mayoría de los destellos de luz detectados por los tubos fotomultiplicadores no corresponderán a productos de desintegración del protón. El número de conos de Cerenkov, así como su orientación, ayudará a distinguir la desintegración del protón de otros sucesos como la colisión de un neutrino creado por rayos cósmicos en la atmósfera terrestre con un protón o un neutrón del detector. La desintegración de un protón en un positrón y un mesón pi neutro, por ejemplo, puede dar lugar a tres conos de luz de Cerenkov formando ciertos ángulos característicos. El positrón generaría un cono; la desintegración del mesón pi en dos fotones, cada uno de los cuales originaría una lluvia de partículas cargadas eléctricamente, sería la causa de los otros dos conos de luz. La interacción de un neutrino provoca una cascada de partículas cargadas eléctricamente y, por ello, se observaría como un único cono de radiación de Cerenkov. (Ilustración realizada por Gabor Kiss.)

Desde 1973 muchos físicos teóricos han contribuido al desarrollo de estas teorías, entre ellos Jogesh C. Pati, de la Universidad de Maryland, y Abdus Salam, del Centro Internacional de Física Teórica de Trieste, Georgi y Sheldon Lee Glashow, de Harvard, Harold Fritsch y Peter Minkowski, del Cal Tech, y Feza Gürsey, Pierre Ramond y Pierre Sikivie, de la Universidad de Yale. Los modelos se conocen en general por el nombre del grupo matemático de simetrías que conectan las distintas fuerzas, como son SU(4)[4], SU(5), SO(10), E_6, E_7 y SU(7). Todos estos modelos incluyen partículas exóticas que, al ser emitidas o absorbidas, transforman un quark en un antiquark, en un leptón o en un antileptón; pueden, por tanto, violar la conservación del número bariónico como ya notaron Pati y Salam en el primer artículo sobre la unificación de las interacciones fuertes y electrodébiles. Además, todos estos modelos (al menos, en algunas de sus versiones) satisfacen las hipótesis generales hechas por Georgi, Quinn y yo mismo; así pues, la escala de masas de las partículas exóticas debe ser del orden de 10^{15} a 10^{16} veces la masa del protón y la vida media del protón debe aproximarse a nuestra estimación de 10^{32} años.

Cálculos más refinados han sido hechos recientemente por muchos físicos teóricos, entre ellos Andrej Buras, John Ellis, Mary K. Gaillard y Dimitris V. Nanopoulos, de la Organización Europea para la Investigación Nuclear (CERN), Terrence J. Goldman y Douglas A. Ross, del Cal Tech, William J. Marciano y Alberto Sirlin, de la Universidad Rockefeller y la Universidad de Nueva York, Cecilia Jarlskog, del CERN, y Francisco Ynduráin, de la Universidad Autónoma de Madrid, y Lawrence Hall, de Harvard. Los nuevos cálculos dan un valor mejorado de unas 10^{15} veces la masa del protón para la escala de masa superpesada y una vida media del protón de alrededor de 10^{31} años. Desgraciadamente, el cálculo de la probabilidad de desintegración del protón es complicado debido a la presencia de fuerzas nucleares que actúan entre los quarks y antiquarks del protón y de los productos de la desintegración; incluso en el caso de que las propiedades de las partículas exóticas superpesadas fueran conocidas con precisión, probablemente sería aún imposible afinar la predicción de la vida media del protón más allá de su orden de magnitud.

Los estudios experimentales de las interacciones débiles han proporcionado ya una limitada verificación del análisis general hecho por Georgi, Quinn y el autor. A nadie sorprendería que dos cualesquiera de las tres gráficas de las constantes de acoplamiento se cruzaran en alguna parte; mas, para que las tres curvas de dependencia con la energía de las tres constantes de acoplamiento se crucen en el mismo punto, es necesario imponer alguna condición a los puntos de partida, esto es, a los valores de las constantes de acoplamiento a baja energía. Nosotros usa-

mos esta condición y, dentro de nuestro análisis general, obtuvimos un valor cercano a 0,2 para cierto parámetro (relacionado con la razón entre g_1 y g_2), que describe la unificación de las interacciones electromagnéticas y débiles. (N. T.: Este parámetro es conocido como el "ángulo de Weinberg".) Los experimentos de interacciones con electrones y con neutrinos dan actualmente para este parámetro un valor entre 0,2 y 0,23. Los valores teórico y experimental son lo suficiente próximos para animarnos a considerar seriamente este análisis, aunque en él se haya hecho una extrapolación sin precedentes en extensión: 13 órdenes de magnitud en la energía.

Una masa del orden de 10^{15} veces la del protón resulta tan grande que la emisión y absorción de tales partículas es casi imposible a energías accesibles experimentalmente y, por tanto, sólo puede producir pequeñísimos efectos en los experimentos factibles. La única esperanza para detectar estos diminutos efectos reside en la posibilidad de que puedan violar leyes de conservación que de otra forma serían exactas y posibilitar, con ello, procesos que de otra forma estarían estrictamente prohibidos. Una de estas leyes de conservación es la del número bariónico, puesta a prueba cuando se busca la desintegración del protón. La única otra ley de conservación conocida que no sea requerida para la coherencia de las teorías de interacciones de partículas y que, por tanto, podría ser violada por efectos de superalta energía, es la del número leptónico, esto es, la conservación del número total de leptones (neutrinos, electrones, muones, etcétera) menos el número de antileptones. La no conservación del número leptónico podría mostrarse en procesos como la desintegración beta doble sin neutrinos: la desintegración de dos protones de un núcleo en dos neutrones más dos positrones. El número bariónico permanecería constante en esta reacción; en cambio, el número leptónico disminuiría en 2 unidades. (La reacción no viola la conservación de la energía, ya que los neutrones del estado final no son libres, sino que tienen una energía de ligadura negativa en el núcleo.) Otra señal de que se viola la ley de conservación de número leptónico sería que el neutrino tuviera masa no nula.

Hay en marcha varios intentos nuevos por determinar la vida media del protón. Algunos experimentos operan ya en la mina de Soudan en Minnesota, el campo de oro de Kolar al sur de la India, el túnel del Mont Blanc entre Francia e Italia y el valle de Baksan en el Cáucaso (Unión Soviética). Antes de finales de año, otros tres experimentos deberían estar tomando datos; son los de la mina de sal de Morton en Ohio, la

LUGAR	ENTRADA EN FUNCIONAMIENTO	PROFUNDIDAD (EN METROS DE AGUA EQUIVALENTES)	MATERIAL OBSERVADO
MINA DE ORO DE HOMESTAKE, DAKOTA DEL SUR	AHORA	4400	150 TONELADAS DE AGUA (SE AUMENTARA A 900 TONELADAS)
CAMPO DE ORO DE KOLAR, INDIA	AHORA	7600	150 TONELADAS DE HIERRO (PLACAS DE 1,3 CENTIMETROS)
VALLE DE BAKSAN, UNION SOVIETICA	AHORA	850	80 TONELADAS DE LIQUIDO DE CENTELLEO
MINA DE SOUDAN, MINNESOTA	AHORA	1800	30 TONELADAS DE HORMIGON TACONITA E HIERRO (SE AUMENTARA A 1000 TONELADAS)
TUNEL DEL MONT BLANC ENTRE FRANCIA E ITALIA	AHORA	4270	30 TONELADAS DE HIERRO Y LIQUIDO DE CENTELLEO (SE AUMENTARA A 200 TONELADAS)
MINA DE SAL DE MORTON, OHIO	1981 (ESTIMADO)	1670	10.000 TONELADAS DE AGUA
MINA DE SILVER KING, UTAH	1981 (ESTIMADO)	1700	1.000 TONELADAS DE AGUA
TUNEL DEL MONT BLANC ENTRE FRANCIA E ITALIA	1981 (ESTIMADO)	5000	INICIALMENTE 150 TONELADAS DE HIERRO (PLACAS DE UN CENTIMETRO DE GROSOR)
TUNEL DEL GRAN SASSO, ITALIA	EN CONSTRUCCION	4000	10.000 TONELADAS DE HIERRO (PLACAS DE TRES MILIMETROS)
JAPON	EN CONSIDERACION	2700	3.400 TONELADAS DE AGUA
TUNEL DE FREJUS ENTRE FRANCIA E ITALIA	EN CONSIDERACION	4500	1.500 TONELADAS DE HIERRO (PLACAS DE TRES A CUATRO MILIMETROS DE GROSOR)
JAPON	EN CONSIDERACION	2700	600 TONELADAS DE HIERRO
MINA DE SAL DE ARTYEMOVSK, UNION SOVIETICA	EN CONSIDERACION	600	100 TONELADAS DE LIQUIDO DE CENTELLEO

TABLA DE EXPERIMENTOS de desintegración del protón. Comprende cinco detectores que están ya en funcionamiento, tres que empezarán a registrar datos a lo largo del año en curso y cinco sujetos a estudio. Los experimentos se diferencian por el tipo y cantidad de material observado. Además,

mina de Silver King en Utah y en otro punto del túnel del Mont Blanc. Otros experimentos se hallan en estudio.

La técnica básica de todos los experimentos consiste en compensar la extrema lentitud de la desintegración del protón observando cuidadosamente una gran masa de material. Cuanto mayor sea ésta, tanto mayor será el número de protones y de neutrones ligados y, por tanto, más alta la probabilidad de observar una desintegración. De esta forma se espera poder detectar la desintegración del protón o del neutrón ligado, aunque su vida media sea bastante más larga que la cota actual de 10^{30} años. Los experimentos se distinguen principalmente por la naturaleza y cantidad del material usado, la naturaleza y disposición de los instrumentos utilizados para la detección de la desintegración del protón en el material y las características del experimento en orden a suprimir las falsas señales originadas por los rayos cósmicos, incluyendo la profundidad a la cual los experimentos son realizados.

Al tener que controlarse una gran masa, en los experimentos debe usarse algún material relativamente barato como el agua, el hormigón o el hierro. En los experimentos de la mina de Soudan, el campo de oro de Kolar y el túnel del Mont Blanc, en los que se usa hierro u hormigón, han de emplearse detectores como los tubos proporcionales o los tubos de "dardos", que pueden detectar directamente las partículas cargadas energéticas que se espera sean emitidas en las desintegraciones protónicas. Las partículas cargadas tienen un recorrido corto en el hierro o el hormigón y, por tanto, los tubos detectores deben distribuirse profusamente por el material.

Por otro lado, en los experimentos de la mina de sal de Morton, la mina de Silver King y la mina de oro de Homestake, en los que se usa un material transparente, como el agua, puede seguirse una estrategia bastante distinta. La energía liberada en la desintegración del protón es suficientemente grande para que un electrón, positrón, muon o mesón pi, emitidos en la desintegración, tengan casi con toda seguridad una velocidad altísima, evidentemente inferior a la velocidad de la luz en el vacío pero superior a la velocidad de la luz en el agua. Cuando una partícula cargada viaja a través de un medio transparente a una velocidad superior a la de la luz en dicho medio, ocurre lo que se llama el efecto Cerenkov. Es como el estruendo sonoro originado por un avión que vuele más deprisa que el sonido en el aire. El efecto Cerenkov es, en cambio, un "estruendo óptico" en el que la partícula emite un cono de luz en lugar de uno de sonido. (El hermoso brillo azul de la luz de Cerenkov fue notado ya en los primeros experimentos de radiactividad por Marie Curie, pero sus propiedades fueron exploradas en detalle por primera vez por Pavel A. Cerenkov en la década de 1930.) El ángulo comprendido entre los rayos de luz de Cerenkov y la trayectoria de la partícula cargada depende de la relación entre la velocidad de la partícula cargada y la de la luz en el medio. En la mayoría de los casos, la luz se emite por partículas que se desplazan a velocidades próximas a la de la luz en el vacío; en cuyo caso, el valor característico del ángulo en agua es de aproximadamente 42 grados.

La observación de un cono de luz de Cerenkov señala que en el medio ha sucedido algo que ha creado una partícula cargada rápida. Además, para una velocidad inicial dada, el grosor del cono y la cantidad de luz emitida dependen sólo de la distancia recorrida por la partícula cargada antes de que su velocidad se reduzca a un valor inferior a la velocidad de la luz en el medio, lo que a su vez depende de su energía inicial. Registrando las posiciones donde se recibe la luz y la intensidad de ésta, podemos deducir, entonces, la energía inicial y la dirección de cada partícula cargada. En determinadas circunstancias, la propia partícula móvil puede desintegrarse emitiendo otras partículas que originan un segundo destello de luz de Cerenkov. Un muon o un antimuón creado en la desintegración de un protón podrían desintegrarse en un electrón que emitiera radiación de Cerenkov. Un mesón pi cargado podría desintegrarse en un muon (o un antimuón) lento que, a su vez, podría desintegrarse en un electrón (o un positrón) rápido que emitiría un cono de

DETECTORES	INSTITUCIONES
144 TUBOS FOTOMULTIPLICADORES EN AGUA	UNIVERSIDAD DE PENNSYLVANIA
CONTADORES PROPORCIONALES DE GAS ENTRE LAS PLACAS	INSTITUTO TATA DE INVESTIGACION FUNDAMENTAL, BOMBAY UNIVERSIDAD DE OSAKA UNIVERSIDAD DE TOKYO
1200 FOTOTUBOS	INSTITUTO PARA INVESTIGACION NUCLEAR, MOSCU
3456 CONTADORES PROPORCIONALES DE GAS EN EL HORMIGON	UNIVERSIDAD DE MINNESOTA LABORATORIO NACIONAL DE ARGONNE
FOTOTUBOS Y TUBOS DE "DARDOS"	INSTITUTO PARA INVESTIGACION NUCLEAR, MOSCU UNIVERSIDAD DE TURIN
2400 TUBOS FOTOMULTIPLICADORES EN AGUA	UNIVERSIDAD DE CALIFORNIA EN IRVINE UNIVERSIDAD DE MICHIGAN LABORATORIO NACIONAL DE BROOKHAVEN
300 TUBOS FOTOMULTIPLICADORES EN AGUA Y PAREDES DE ESPEJO	UNIVERSIDAD DE HARVARD UNIVERSIDAD DE PURDUE UNIVERSIDAD DE WISCONSIN
TUBOS DE "DARDOS" ENTRE LAS PLACAS	LABORATORIO DEL SINCROTRON NACIONAL ITALIANO EN FRASCATI UNIVERSIDAD DE MILAN UNIVERSIDAD DE TURIN
CAMARAS DE DESTELLOS ACCIONADAS POR TUBOS DE "DARDOS"	LABORATORIO DEL SINCROTRON NACIONAL ITALIANO EN FRASCATI UNIVERSIDAD DE MILAN UNIVERSIDAD DE TURIN UNIVERSIDAD DE ROMA
1056 TUBOS FOTOMULTIPLICADORES DE 50 CENTIMETROS	UNIVERSIDAD DE TOKYO LABORATORIO NACIONAL JAPONES DE FISICA DE ALTAS ENERGIAS UNIVERSIDAD DE TSUKUBA
TUBOS DE DESTELLOS Y TUBOS DE "DARDOS"	ECOLE POLYTECHNIQUE UNIVERSIDAD DE PARIS EN ORSAY ESCUELA DE MINAS DE SACLAY
TUBOS DE DESTELLOS DE NEON	UNIVERSIDAD DE OSAKA UNIVERSIDAD DE TOKYO LABORATORIO NACIONAL JAPONES DE FISICA DE ALTAS ENERGIAS
128 TUBOS FOTOMULTIPLICADORES	INSTITUTO PARA INVESTIGACION NUCLEAR, MOSCU

se realizan a distintas profundidades bajo tierra. La profundidad viene expresada en términos de la profundidad equivalente de agua que proporcionaría la misma protección respecto a los rayos cósmicos.

luz. Un mesón pi neutro podría desintegrarse en dos fotones, cada uno de los cuales originaría una lluvia de partículas cargadas acompañadas de luz de Cerenkov. Así, el uso de la luz de Cerenkov para detectar desintegraciones de protones ofrece una alternativa a la observación directa de partículas cargadas como medio de reconstrucción del proceso de desintegración y de verificación de que se trata realmente de una desintegración de un protón.

Una ventaja de la luz de Cerenkov con respecto a los demás sistemas de detección de la desintegración del protón es que la luz puede recorrer, en el agua, distancias superiores a las recorridas por las propias partículas cargadas. De esta forma, para un determinado volumen de material a observar se necesitan menos detectores que en los experimentos en los que se usan materiales opacos, como el hierro o el hormigón. Sin menospreciar, además, su menor coste. Por otra parte, los detectores Cerenkov para agua sólo son sensibles a partículas cargadas que se muevan más velozmente que la luz en el agua. La relativamente baja densidad del agua exige la excavación de cavidades mayores bajo tierra para albergar una determinada masa de material a observar y este agua, por otro lado, debe conservarse muy pura para mantener su transparencia a la luz de Cerenkov.

¿Qué ritmos de desintegración del protón podrían detectarse en estos experimentos? Tomemos como ejemplo el experimento de la mina de sal de Morton, el proyectado con la mayor cantidad de masa a observar. De sus 10.000 toneladas de agua, se usará una capa exterior de unas 5000 toneladas para controlar la señal de fondo originada por los rayos cósmicos. Las restantes 5000 toneladas de agua contienen 10^{33} protones (y neutrones ligados). Si la vida media del protón es de 10^{31} años como indican las versiones refinadas del análisis hecho por Georgi, Quinn y el autor, deberíamos tener unos 300 sucesos de desintegración del protón por año. Algunos años de observación nos proporcionarían aún unos pocos sucesos de desintegración del protón si su vida media alcanzase los 10^{33} años, pero con un ritmo de desintegración tan lento, los experimentos empezarían a contaminarse de un fondo imposible de substraer originado por falsos sucesos provocados por neutrinos de rayos cósmicos. Sería muy difícil mejorar los experimentos en tales circunstancias.

¿Qué habremos aprendido si se descubre la desintegración del protón? Una conclusión inmediata será que el número bariónico no se conserva, lo que confirmará, una vez más, la opinión en auge de que todas las cantidades conservadas poseen un significado dinámico similar al de la carga eléctrica. Además, si la desintegración del protón se descubre en un futuro próximo, la vida media deberá estar comprendida entre 10^{30} y 10^{33} años, lo que dará cierta credibilidad a las hipótesis generales sobre la unificación de las fuerzas electrodébiles y fuertes usadas por Georgi, Quinn y el autor. Existen, sin embargo, muchísimas teorías que satisfacen estas hipótesis generales, como son algunas versiones de los modelos $SU(4)^4$, SU(5), SO(10), etcétera, anteriomente mencionados. Será difícil discernir cuál de estas teorías (si es alguna de ellas) describe la física a muy alta energía.

De una cosa podemos estar seguros. Si se llega al descubrimiento de la desintegración del protón, se dedicarán nuevos grandes medios a su estudio y pronto habrá una segunda generación de experimentos en los que el esfuerzo no se centrará ya en averiguar si el protón se desintegra, sino en cómo se desintegra: cuáles son las probabilidades de los distintos modos de desintegración.

Como preparación a este esfuerzo, un elevado número de físicos teóricos han venido explorando los modos probables de desintegración del protón. (Los comentarios que siguen están basados en trabajos independientes realizados por Wilczek y por Anthony Zee y el autor.) Es interesante el que uno pueda ir bastante lejos en estos análisis sin hacer hipótesis sobre la unificación de las interacciones fuertes y electrodébiles. Todo lo que se necesita son las conocidas leyes de conservación de la carga, color y restantes propiedades y la hipótesis de que las partículas exóticas responsables de la desintegración del protón son muy pesadas, como seguramente deben ser para explicar la larga vida media del protón. Aunque la emisión y reabsorción de estas partículas puede producir muchísimos modos de desintegración del protón, los más complicados están reprimidos más fuertemente que los más simples, debido a la gran masa de las partículas exóticas. Salvo que intervenga alguna circunstancia especial, los modos dominantes de desintegración serán, en general, aquellos en los que el protón o neutrón ligado se desintegre en un antileptón (positrón, antimuón o antineutrino) y algunos mesones, y no aquellos que lo hagan en un leptón (electrón, muon o neutrino) y algunos mesones. Se puede ir más lejos y hacer predicciones acerca de relaciones entre probabilidades de desintegración. Por ejemplo, un neutrón se desintegra el doble de rápido que un protón en un positrón y un único mesón pi o mesón rho. El protón se desintegra más rápidamente en un positrón y mesones que un neutrón en un antineutrino y mesones. El neutrón se desintegra más rápidamente en un positrón y mesones que lo hace el protón en un antineutrino y mesones.

No podemos tener seguridad de que estas predicciones las confirme la experiencia. Si tal no ocurre, ha de haber partículas exóticas bastante más ligeras que 10^{14} veces la masa del protón, que produzcan modos de desintegración más complicados. Por ejemplo, la desintegración de un protón o un neutrón ligado en un electrón y mesones más que en un positrón y mesones podría producirse a un ritmo observable si existiesen partículas exóticas que no pesaran más de 10^{10} veces la masa del protón. Una desintegración en tres neutrinos o tres electrones (o alguna otra combinación de tres leptones) podría observarse si hubiera partículas exóticas que no pesaran más de 10^4 veces la masa del protón. Sin embargo, partículas exóticas relativamente ligeras como éstas deberían poseer propiedades especiales para evitar que se produjera la desintegración "ordinaria" del protón (en un mesón y un positrón o un antineutrino) en una proporción mucho más alta.

La verificación de estas predicciones, esto es, descubrir que el protón se desintegra en mesones y un positrón o un antineutrino, y con las relaciones de ritmos de desintegración anteriormente mencionadas confirmaría que la desintegración del protón se debe realmente a partículas exóticas de masa superior a 10^{10} veces la masa del protón, pero no apuntaría hacia ninguna teoría subyacente específica. Para este fin sería necesario explorar detalles más finos del proceso de desintegración. (Por ejemplo, la determinación de la orientación del spin del positrón o antimuón producido en la desintegración de un protón puede usarse para diagnosticar el spin de las partículas exóticas superpesadas cuya emisión y reabsorción ha producido la desintegración.) Si llega a descubrirse la desintegración del protón, se situará como un triunfo del ingenio experimental y una clave sin paralelo de la física de altas energías, pero presentará nuevas tareas a realizar para entender dicha desintegración.

Antibióticos beta-lactámicos

Las penicilinas y las cefalosporinas se caracterizan, químicamente, por poseer un anillo beta-lactámico. Se ha modificado su estructura molecular para elevar su eficacia y combatir las bacterias resistentes

E. P. Abraham

Han transcurrido casi 40 años desde la introducción de la penicilina en clínica médica. Poco antes se había demostrado la eficacia de las sulfamidas y otras moléculas de pequeño tamaño en el tratamiento de la fiebre puerperal y otras infecciones generalizadas graves. Muchos de los médicos actuales no conocieron la época en la que poco podía hacerse para tratar enfermos con infecciones graves, cuando la endocarditis bacteriana era casi invariablmente fatal, la meningitis meningocócica apenas dejaba supervivientes, y éstos gravemente incapacitados, y la neumonía neumocócica era conocida como "la amiga del anciano" por ser una causa benigna de muerte. Merced a su notable actividad frente a distintas bacteria patógenas y su casi nula toxicidad para el hombre, la penicilina parecía poseer propiedades milagrosas y abrió una nueva era en el tratamiento químico de la enfermedad.

Después del éxito de la penicilina, una vasta investigación de carácter internacional permitió el descubrimiento de un sinnúmero de otras sustancias naturales con propiedades antibacterianas. Muchas de ellas resultaron tóxicas, y se desecharon muy pronto. Otras, como la bacitracina, tuvieron sólo una importancia clínica limitada. Entre las que han perdurado, cabe mencionar la estreptomicina (y otros aminoglicósidos, como la gentamicina) las eritromicinas y la familia de las tetraciclinas. A las primeras penicilinas sucedieron otras de probada eficacia contra cepas que habían desarrollado resistencia frente a la penicilina. Las bases de la resistencia frente a los antibióticos y de los mecanismos implicados en la actividad de los antibióticos constituyen hoy ramas muy activas de la biología celular y la genética. La familia de la penicilina se ha visto ampliada con el advenimiento de las cefalosporinas, químicamente relacionadas con aquéllas, descubiertas en Oxford y que representan del 15 al 20 por ciento de todos los medicamentos prescritos en los hospitales norteamericanos.

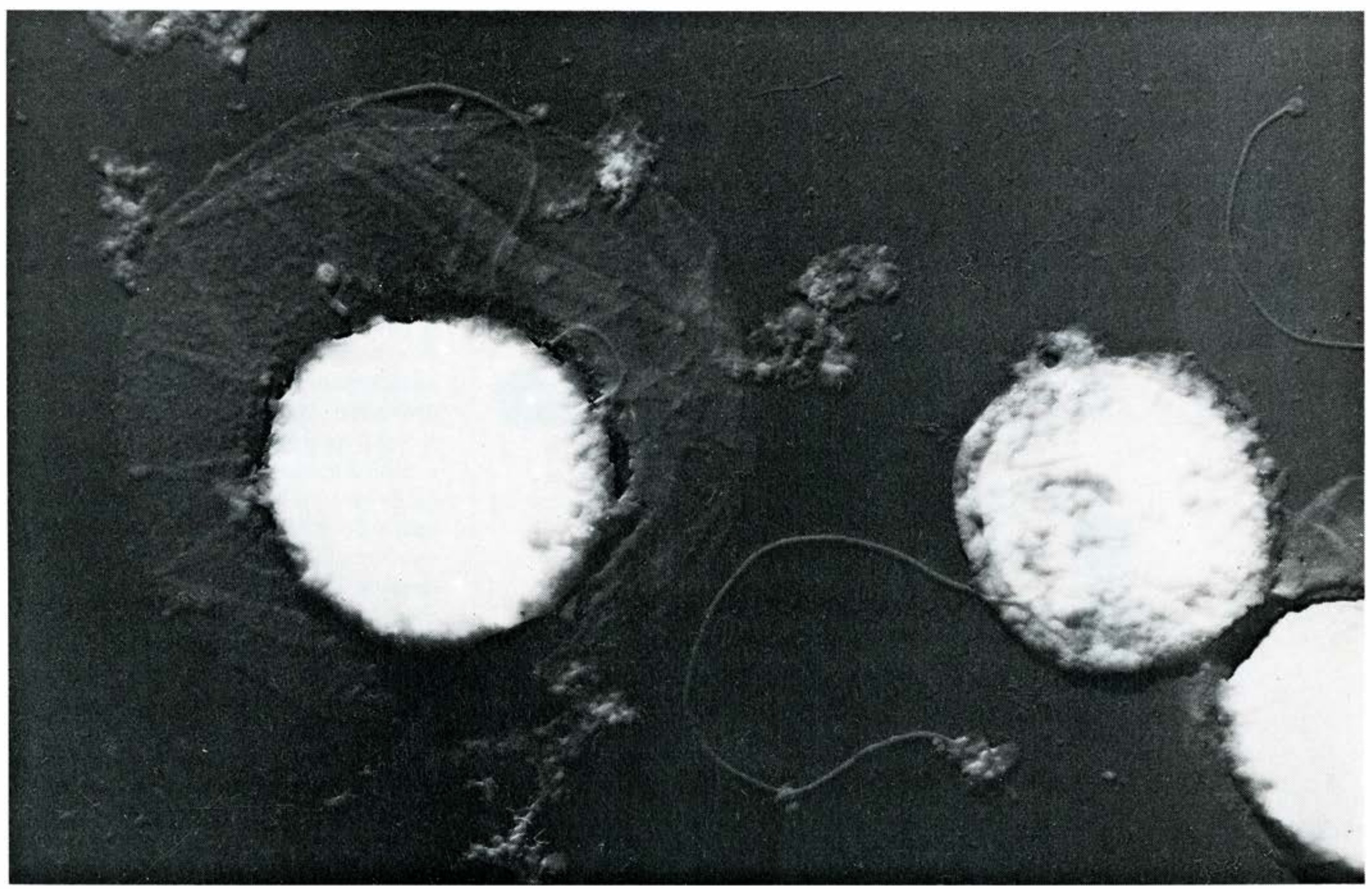
El hecho crucial que abrió la nueva era de los antibióticos en medicina fue la decisión de Howard W. Florey y Ernst B. Chain, en 1938, de investigar sistemáticamente la actividad antibacteriana, conocida desde hacía tiempo, de una serie de sustancias producidas por microorganismos. Estas investigaciones las llevaron a cabo en la Escuela de Patología Sir William Dunn de Oxford. El interés de Florey y Chain era puramente científico y no pretendían que sus investigaciones condujeran al hallazgo de sustancias de utilidad clínica. Por fortuna, la penicilina se encontraba entre las tres sustancias a estudiar en primer lugar.

La penicilina no constituía, de suyo, ninguna novedad. En 1929, Alexander Fleming denominó así un filtrado activo del medio de cultivo del hongo *Penicillium notatum* obtenido en el hospital St. Mary de Londres. Fleming estudiaba variantes de estafilococo, bacteria que puede producir furúnculos, abscesos y, algunas veces, infecciones generalizadas fatales. El hongo se posó por casualidad en una placa que contenía agar en la que Flemig había sembrado bacterias antes de unas vacaciones veraniegas. Un período de clima frío permitió que creciera el hongo y segregara penicilina antes de que se produjera el crecimiento bacteriano completo, que requiere temperaturas más elevadas. A su vuelta al laboratorio, Fleming observó que las colonias de bacterias que habían crecido en la vecindad del *Penicillium* estaban parcialmente lisadas o disueltas. La lisis no habría ocurrido con un clima más cálido, pues los estafilococos se hubieran desarrollado en primer lugar. La penicilina determina la lisis de las bacterias en crecimiento, pero no de aquellas cuyo desarrollo ha llegado a término.

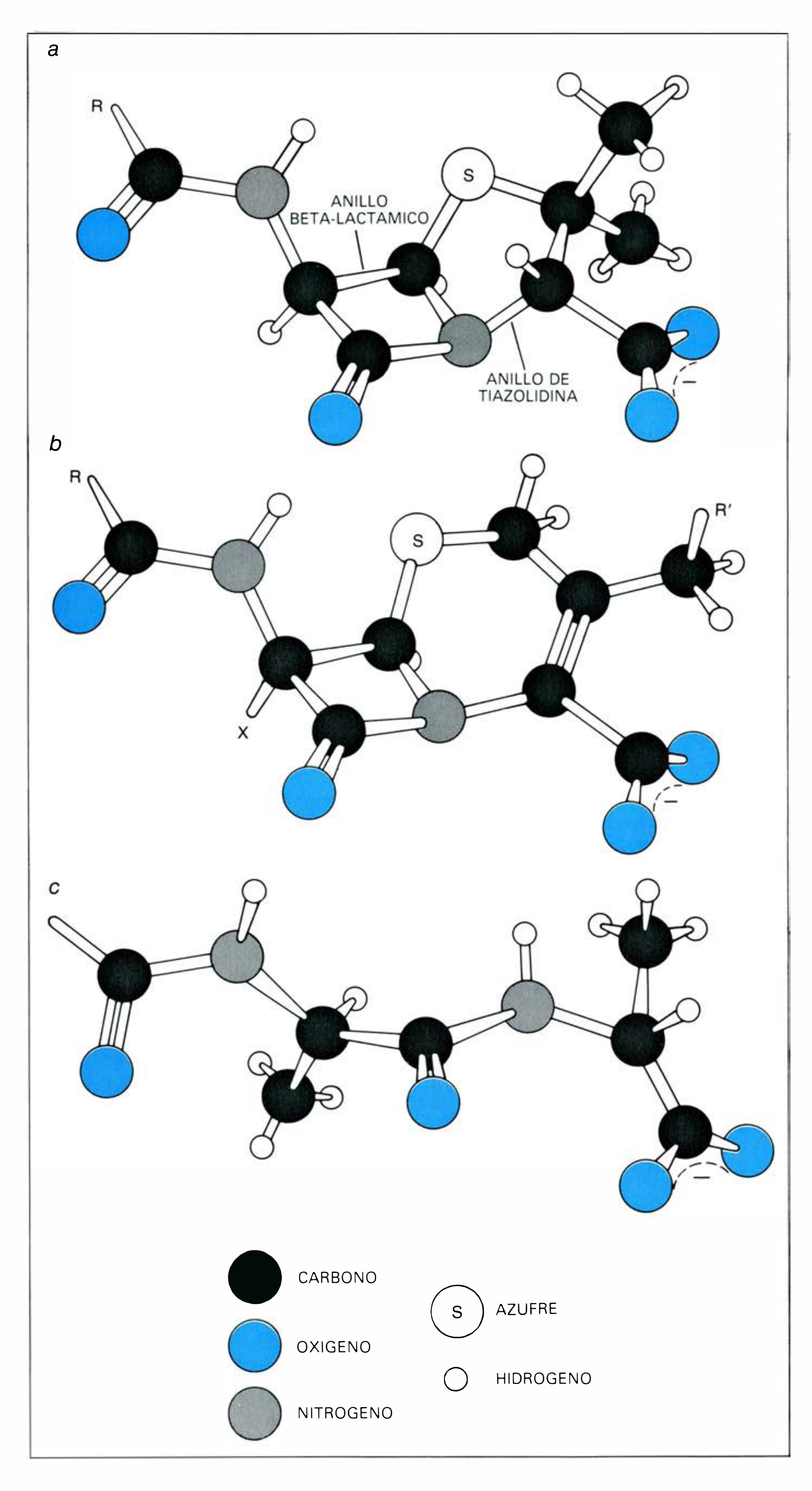
Fleming no fue el primero en observar que los hongos del género *Penicillium* provocaban la inhibición del crecimiento bacteriano en su vecindad. De observaciones similares dieron cuenta John Burdon Sanderson en 1870, Joseph Lister, un año después, William Roberts en 1874, John Tyndall en 1876 y el estudiante de medicina francés Ernest Duchesne, en 1896. Nunca se sabrá si estos investigadores vieron en realidad los efectos de la penicilina, puesto que distintas especies de *Penicillium* segregan un número considerable de otras sustancias antibacterianas.

En todo caso, Fleming fue un agudo observador, que se interesó ulteriormente por el fenómeno contemplado. Comprobó que la penicilina se mostraba muy activa frente a bacterias grampositivas, tales como el estreptococo, estafilococo y pneumococo, pero no frente a bacilos patógenos gram-negativos, como la *Salmonella typhi*, *Escherichia coli* y *Pseudomonas aeruginosa*. La división de las bacterias en gram-positivas y gram-negativas obedece a su diferente respuesta a una técnica de tinción desarrollada por Hans Christian Joachim Gram, en 1884. Los microbiólogos atribuyen hoy esas diferencias a las características propias de la pared celular de las distintas especies.

Fleming demostró también que el caldo de cultivo que contenía penicilina no resultaba más tóxico para los glóbulos blancos de la sangre y para el conejo que el medio de cultivo ordinario. Fleming no valoró, sin embargo, la magnitud de su descubrimiento. Si bien aplicó la penicilina en varias ocasiones como un antiséptico local, en 1940 escribió que su utilidad no justificaba el trabajo de prepararla. Su mayor interés por la sustancia consistía en utilizarla como medio de aislamiento de ciertos organismos en cultivos mixtos, en virtud de su acción antibacteriana selectiva. Nunca pensó que la penicilina podría introducirse en la corriente sanguínea para curar infecciones. Un largo decenio después de su hallazgo, la peni-



EFECTO DE LA PENICILINA, mostrado en micrografías electrónicas realizadas por Milton R. J. Salton, de la Facultad de Medicina de la Universidad de Nueva York. La micrografía superior corresponde a bacterias normales de la especie *Vibrio metchnikovi*. La micrografía inferior muestra los "esferoplastos", que se desarrollan cuando las bacterias crecen en presencia de penicilina. Alrededor del esferoplasto, a la izquierda, se puede observar la tenue pared desprendida de la bacteria. La penicilina bloquea la síntesis de la pared celular de las bacterias sensibles cuando éstas se encuentran aún en período de crecimiento. Lo único que evita la desintegración de los esferoplastos es la frágil membrana citoplasmática, apenas apreciable en la micrografía.



ESTRUCTURA FUNDAMENTAL de las penicilinas y cefalosporinas; consiste en un anillo beta-lactámico. En la penicilina (*a*), el anillo beta-lactámico de 4 átomos está unido al de la tiazolidina, de 5, que consta de un átomo de azufre, uno de nitrógeno y tres de carbono. La cadena lateral *R* es variable; en la página siguiente se han representado varios tipos de cadenas laterales. En las cefalosporinas, químicamente relacionadas (*b*), *R* y *R'* son variables y *X* puede ser un hidrógeno o un grupo metoxil (-OCH_3). Una idea de cómo pueden los antibióticos beta-lactámicos ejercer su acción sobre la pared bacteriana se muestra en la estructura (*c*) de la cadena en crecimiento D-alanil-D-alanina del peptidoglicano de una pared bacteriana. La conformación de esta cadena es similar, en parte, a la de la molécula de penicilina.

cilina era, todavía, apenas algo más que mera curiosidad.

La situación cambió espectacularmente cuando Florey, Chain y un pequeño grupo de colaboradores de Oxford lograron una purificación parcial de la penicilina: demostraron, primero en el ratón y luego en seres humanos, que encerraba extraordinarias propiedades terapéuticas. En 1941, un pequeño ensayo clínico, limitado por las mínimas cantidades de penicilina de que se disponía (se economizaba el producto hasta el extremo de recuperarlo de la orina de los pacientes), proporcionó los primeros entre una serie de resultados que permitieron describir a Florey, en 1949, los resultados de la penicilina como "tan satisfactorios en verdad, que a veces parecen increíbles".

La demostración de la utilidad clínica de la penicilina estimuló a los gobiernos de Gran Bretaña y Estados Unidos a apoyar los enormes esfuerzos de las compañías farmacéuticas en la producción de suficiente cantidad de penicilina para tratar a las víctimas de la segunda guerra mundial. Los rápidos progresos alcanzados se debieron, en gran medida, a la contribución norteamericana, cifrada en la introducción de un nuevo medio de cultivo, la fermentación profunda en lugar de superficial y el elevado rendimiento en las cepas mutantes de *Penicillium chrysogenum*. Estos tres avances permitieron superar lo que al principio parecía un obstáculo insalvable. La cantidad de penicilina que hoy se produce por litro de cultivo multiplica, miles de veces, la que se obtenía en la década de 1940 en Oxford.

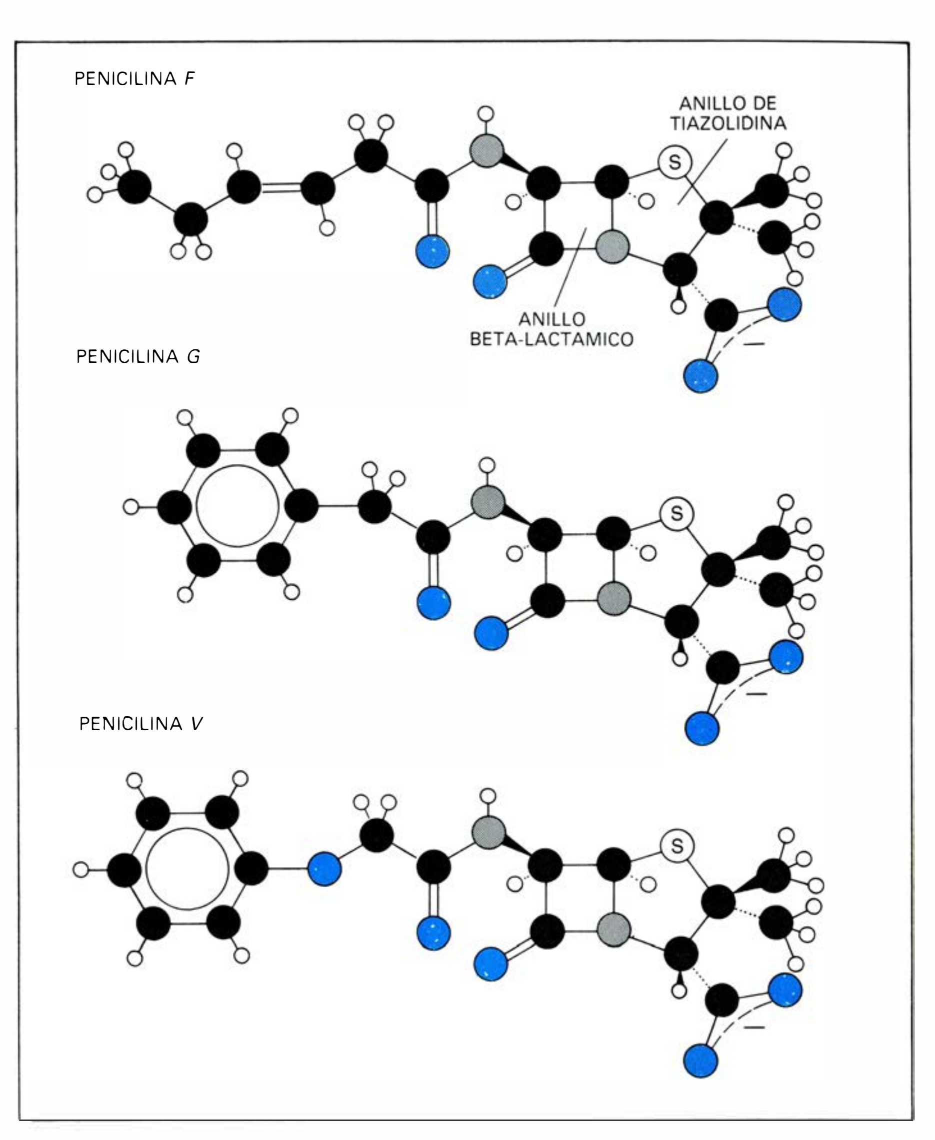
Los años de la guerra supusieron también una intensa investigación angloamericana, protegida por ambos gobiernos, para producir penicilina por síntesis química. El éxito dependía del establecimiento de la estructura química de la molécula. En 1943, Chain y el autor de este artículo propusieron como estructura de la penicilina un anillo beta-lactámico de 4 átomos unido a otro de tiazolidina, de 5 átomos; constaba éste de un átomo de azufre, uno de nitrógeno y tres de carbono. La configuración beta-lactámica recibiría, más tarde, un fuerte apoyo por parte de Robert B. Woodward, de la Universidad de Harvard, pero no se aceptó definitivamente hasta 1945, cuando los análisis cristalográficos de rayos X, realizados por Dorothy C. Hodgkin y Barbara W. Low, en Oxford, demostraron con nitidez las posiciones relativas de los átomos en la molécula.

Fracasaron los intentos por lograr la

síntesis de la inestable molécula de penicilina. El éxito llegaría el año 1958, cuando John C. Sheehan y sus colaboradores, del Instituto de Tecnología de Massachusetts, sintetizaron la penicilina con ayuda de un nuevo reactivo que les permitió cerrar el anillo beta-lactámico en condiciones químicas suaves. Sin embargo, la síntesis química no podía competir económicamente con la fermentación en la producción de penicilina. La semisíntesis a partir de un producto de fermentación condujo a la obtención de sustancias de gran valor terapéutico.

Al comienzo de la década de 1950, la utilización de la penicilina abrió una nueva era en la historia de la medicina. Su actividad altamente selectiva frente a las células bacterianas le confirió el extraordinario valor que todavía posee. No obstante, la penicilina de que se disponía en aquellos momentos tenía sus limitaciones; en particular, había bacterias resistentes a la misma. Ya en 1940, trabajando con Chain, descubrimos un enzima de *E. coli* que destruía la penicilina de Oxford. Sugerimos entonces que el enzima, al que habíamos bautizado penicilinasa, intervenía en la resistencia de las bacterias que lo producían.

Incluso antes de descubrirse las características estructurales del anillo típico de la penicilina, se sabía que la molécula producida en Oxford difería de la obtenida en los laboratorios de los Estados Unidos. Las dos penicilinas, que se denominaron, respectivamente, penicilina *F* y penicilina *G*, tenían el mismo núcleo molecular, pero diferían en las cadenas laterales. La penicilina *G*, conocida también por bencilpenicilinasa, fue la primera que alcanzó un uso generalizado en medicina, y aún mantiene su valor terapéutico. Se obtuvo en los Estados Unidos al adicionar un extracto de cereal que contenía ácido fenilacético ($C_6H_5CH_2COOH$) al medio de fermentación. La cadena lateral deriva del citado ácido.



CADENAS LATERALES de tres penicilinas, representadas en cada caso a la izquierda de la estructura fundamental beta-lactámica. La penicilina *F* es la 2-pentenilpenicilina. La *G* es la bencilpenicilina y la *V* es la primera penicilina oral. Las cadenas representadas son las laterales *R* de la molécula fundamental.

Andando el tiempo, se obtuvieron distintas penicilinas con cadenas laterales químicamente relacionadas, a partir de fermentaciones de *Penicillium chrysogenum* a las que se agregaban los precursores adecuados. Ninguna de estas penicilinas resultó marcadamente superior a la bencilpenicilina en su rango de actividad. En 1954, sin embargo, se descubrió que la fenoximetilpenicilina, a diferencia de la bencilpenicilina, era estable en el medio ácido gástrico, por lo que podía utilizarse por vía oral.

Por aquel entonces no había datos que permitieran suponer la posible obtención de compuestos de la familia de la penicilina exentos de las limitaciones de la bencilpenicilina. De hecho, había indicios de que el valor terapéutico de la penicilina disminuiría ante la progresiva presencia, en los hospitales, de estafilococos resistentes a la bencilpenicilina. Tal resistencia la confería su capacidad de sintetizar una penicilinasa. Afortunadamente, varios hechos imprevisibles condujeron a la obtención de nuevas penicilinas, de cefalosporinas y otras sustancias con el anillo beta-lactámico, que retenían la mayoría de las propiedades terapéuticas de la bencilpenicilina y, además, eran eficaces frente a muchas bacterias resistentes a la bencilpenicilina. Tales compuestos, junto a las primeras penicilinas, constituyen el grupo actual de antibióticos beta-lactámicos. Al mismo tiempo, las investigaciones permitieron descifrar el modo de acción de los antibióticos beta-lactámicos, las razones de su diferente actividad y los factores que crean la resistencia bacteriana y permiten la propagación de los microorganismos.

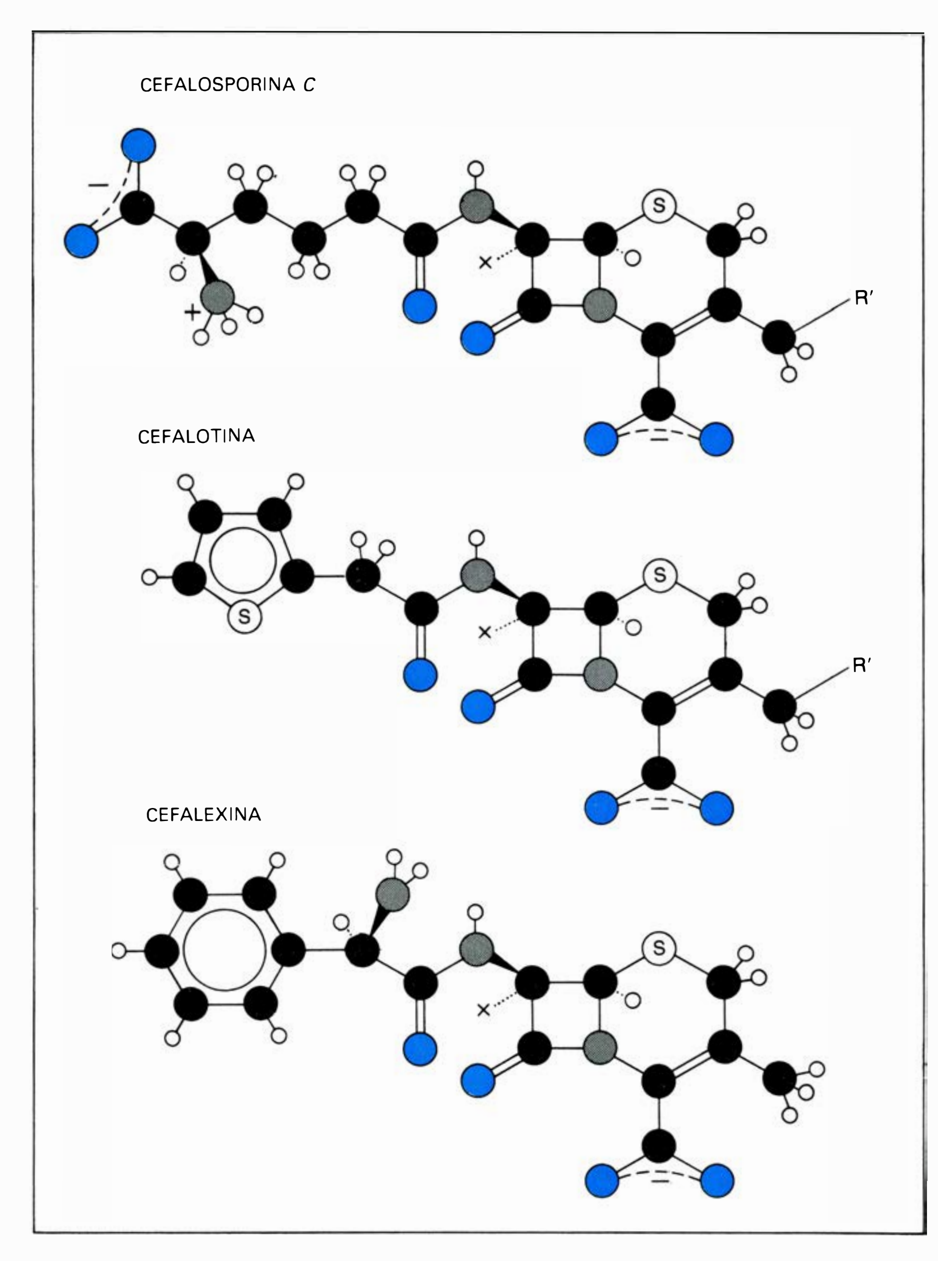
Partiendo de la observación de que la penicilina destruía únicamente bacterias en crecimiento, Joshua Lederberg, a la sazón en la Universidad de Wisconsin, mostró que las células de *E. coli* en forma de bastoncito, cuando crecían en un medio hipertónico (con una presión osmótica anormalmente alta), daban lugar a células esféricas si se añadía penicilina al medio de cultivo. Estos "esferoplastos" tenían una frágil membrana citoplasmática, pero carecían de la mayor parte de la pared celular propia de sus progenitores. Por cuya razón, las células estallaban cuando bajaba, por dilución, la presión osmótica del medio. Lederberg concluyó que la penicilina bloqueaba la síntesis de la pared bacteriana.

La naturaleza de tal interferencia y

sus consecuencias han podido conocerse gracias a los datos obtenidos sobre la estructura molecular de la pared bacteriana por Milton R. J. Salton, de la Facultad de Medicina de la Universidad de Nueva York, por Jack L. Strominger, de Harvard, y por James Baddiley, de la Universidad de Newcastle en Tyne, entre otros. Tanto en bacterias gram-positivas, como gram-negativas, la membrana citoplasmática está cubierta por una capa de peptidoglicanos, conocida también como mucopéptido o mureína. Los peptidoglicanos forman una estructura tridimensional cuyo esqueleto consta de unidades alternantes de dos aminoazúcares, la N-acetilglucosamina (NAG) y el ácido N-acetilmurámico (NAM).

Las unidades NAM, que se encuentran únicamente en las paredes bacterianas o de microorganismos próximos, están unidas a cadenas peptídicas (cadenas cortas de aminoácidos) que contienen tanto aminoácidos D (dextro) como L (levo). Muchas de estas cadenas peptídicas se entrecruzan. La estructura resultante proporciona a la célula una pared rígida y le confiere su forma característica. La capa de peptidoglicanos de las bacterias gram-negativas es más delgada y tiene menos entrecruzamientos que la propia de las bacterias gram-positivas. Está también rodeada por una membrana externa con fosfolípidos, lipopolisacáridos y proteínas con una parte interior hidrofóbica (repelente del agua).

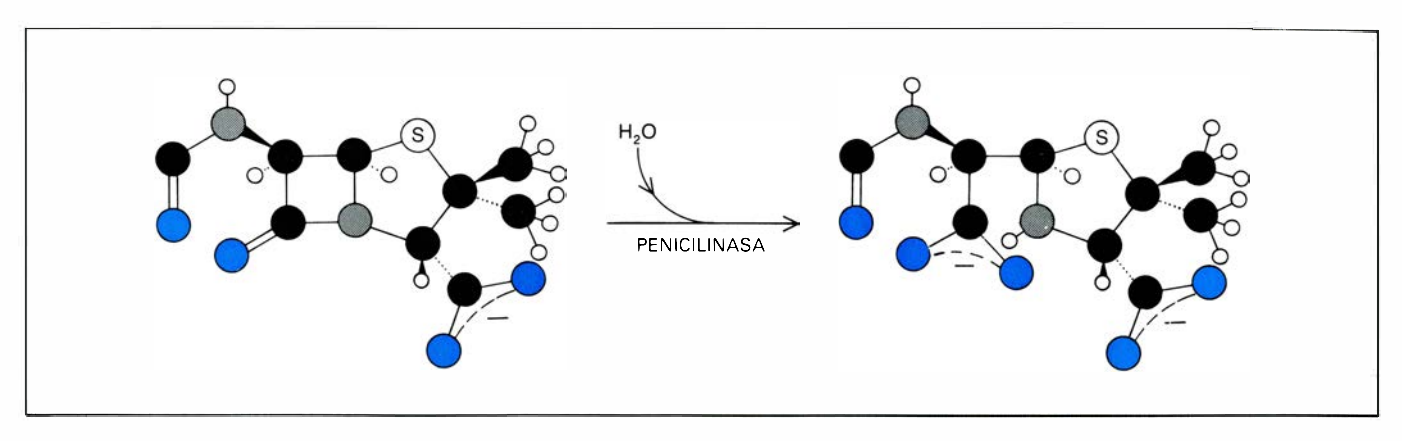
Los estudios llevados a cabo por Strominger, James T. Park, de la Facultad de Medicina de la Universidad de Tufts, y por J. M. Ghuysen, de la Universidad de Lieja, han demostrado que las membranas citoplasmáticas de las bacterias y de hongos próximos de la familia *Streptomyces* poseen enzimas que resultan inactivados por la penicilina y otros antibióticos beta-lactámicos. Los enzimas vulnerables son, bien transpeptidasas (que entrecruzan una cadena peptídica de peptidoglicano con otra por desplazamiento de una D-alanina terminal) o D-carboxipeptidasas (que eliminan una D-alanina terminal por hidrólisis sin que tenga lugar entrecruzamiento) o enzimas que pueden realizar ambas acciones. Se considera que la inhibición del entrecruzamiento constituye la reacción a través de la cual los antibióticos beta-lactámicos ponen en marcha una serie de mecanismos que conducen a la muerte, y a menudo a la lisis, de la célula bacteriana. Puesto que la síntesis de peptidoglicanos no tiene lugar en células humanas y de otros mamíferos, no es sorprendente que las sustancias seleccionadas en el curso de la evolución para inhibir dicha síntesis resulten relativamente inocuas para el organismo humano.



ESTRUCTURA DE LA CADENA LATERAL de tres cefalosporinas, representada como en el caso de las penicilinas de la ilustración de la página anterior. La cefalosporina *C* fue la primera descubierta de la nueva familia de penicilinas derivadas del hongo *Cephalosporium* (en lugar del tradicional *Penicillium*).

Los mecanismos implicados directamente en la destrucción y lisis de las bacterias no se conocen con exactitud. Las células bacterianas poseen enzimas autolíticos capaces de destruir los peptidoglicanos, además de los enzimas necesarios para su síntesis. Alexander Tomasz, de la Universidad Rockefeller, observó que los neumococos que portaban una mutación causante de un defecto de su sistema autolítico se libraban de la lisis por penicilina, si bien se inhibía su crecimiento. Tales organismos se han descrito como tolerantes. La cuestión que se plantea es si el sistema autolítico de cepas normales se dispara a raíz de la acción de la penicilina y es responsable de la lisis producida por el antibiótico. No se sabe en qué medida opera tal mecanismo, pero no hay duda de que los antibióticos betalactámicos actúan a través de un sistema de considerable complejidad.

Investigaciones recientes han puesto de manifiesto que la membrana citoplasmática de las células bacterianas posee múltiples proteínas que se unen específicamente a los antibióticos betalactámicos; entre ellas se encuentran los enzimas con actividad D-alanina carboxipeptidasa o transpeptidasa, o ambas. Las proteínas con capacidad de unión de la penicilina (PBP) pueden ponerse de manifiesto solubilizando la membrana con un detergente, después de haber incubado las células con penicilina marcada con carbono 14, y observando las proteínas marcadas, previa



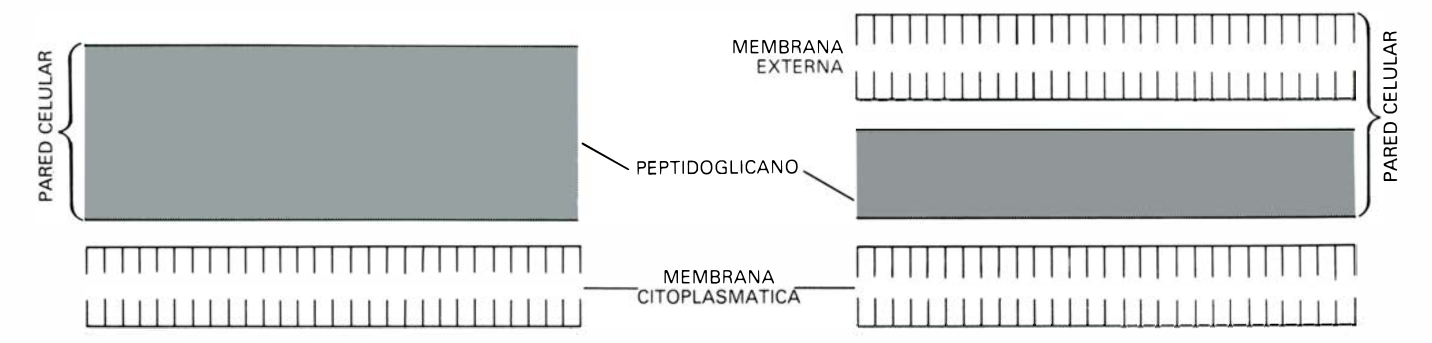
LAS BACTERIAS RESISTENTES representaron un problema en el uso clínico de la penicilina. En 1940, Ernst B. Chain y el autor descubrieron, en la Universidad de Oxford, un enzima de la bacteria *Escherichia coli* que destruía la penicilina. Lo denominaron penicilinasa. Este enzima y otros similares actúan rompiendo, por hidrólisis, el enlace amida del anillo beta-lactámico de la penicilina; se forma ácido peniciloico, que no tiene actividad antibiótica.

separación por electroforesis en gel de agar. Diferentes antibióticos beta-lactámicos poseen afinidades distintas para una o más PBP y provocan diversos cambios morfológicos en las bacterias. Cuando actúan sobre el bacilo *E. coli* a concentraciones suficientemente altas, la mayoría desencadena la formación de protoplastos y la lisis celular, mientras que a concentraciones más bajas algunos inducen la formación de células esféricas osmóticamente estables o de filamentos de células, en los que se han inhibido la división celular.

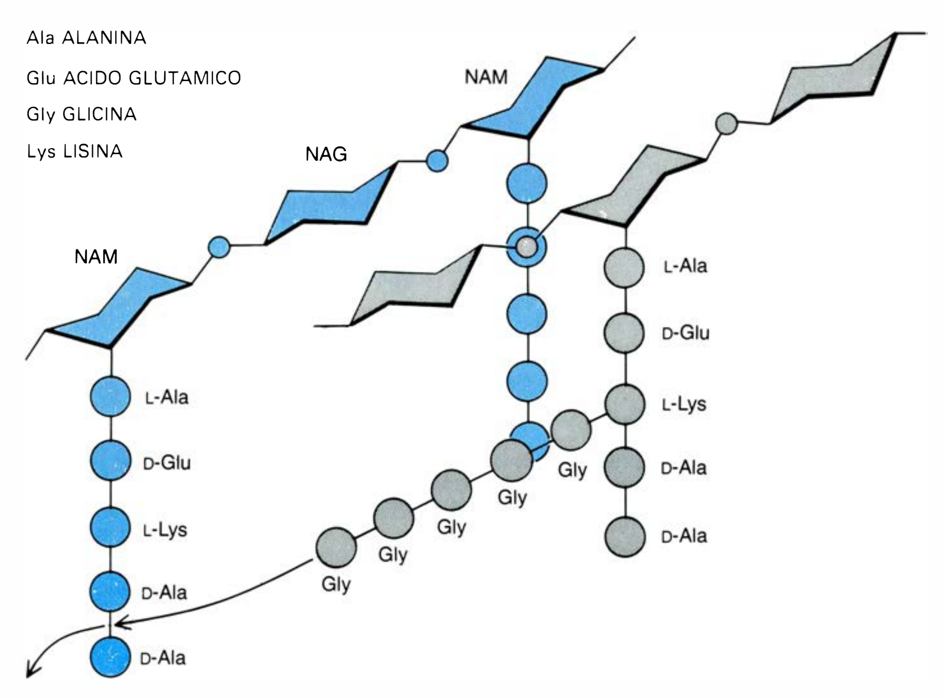
El crecimiento de mutantes de *E. coli* en los que algunas PBP son estables a 30 grados Celsius, pero inestables a 42 grados, se acompaña a veces de cambios morfológicos a la temperatura más alta, que mimetizan a los causados por los antibióticos beta-lactámicos a la temperatura más baja. Un estudio de estos mutantes llevado a cabo por Brian G. Spratt, de la Universidad de Sussex, mostró que una PBP denominada 1B estaba relacionada con la elongación celular, la PBP2 con la forma de la célula y, la PBP3, con la división celular. La inhibición de cualquiera de estas PBP es letal, mientras que la de las demás no lo es, aunque se encuentran en mayor cantidad.

Cualquier disminución de la afinidad de los antibióticos beta-lactámicos por PBP esenciales o la disminución de su capacidad para alcanzar el punto donde se encuentran dichos enzimas quedarán reflejadas en su actividad antibacteriana. Se ha aislado un mutante de *E. coli* sensible a la temperatura que muestra una mayor resistencia a dos antibióticos beta-lactámicos (mecilinam y tienamicina) y una disminución de la afinidad de la PBP2 para tales sustancias. Se desconoce actualmente si las resistencias adquiridas de esta forma, por mutación, tienen importancia clínica. Ambos antibióticos ejercen sus efectos letales sobre *E. coli* uniéndose, casi exclusivamente, a PBP2. Se requerirían mutaciones que afecten a la síntesis de varias PBP para conferir resistencia frente a otros antibióticos beta-lactámicos. Además, los cambios producidos en las PBP deberían ser lo bastante específicos como para inhibir la unión de los antibióticos sin obstruir la unión de los sustratos naturales requeridos para la síntesis de los peptidoglicanos, esenciales para la supervivencia de la célula.

Las bacterias gram-negativas tienden a ser intrínsecamente más resistentes frente a muchos antibióticos beta-lactámicos que las gram-positivas. Los resultados de las investigaciones sobre las propiedades de la membrana externa de las bacterias gram-negativas, llevados a cabo por Hiroshi Nikaido, de la Universidad de California en Berkeley, y por otros apoyan la hipótesis de que dicha membrana contiene proteínas (denominadas porinas) que proporcionan canales a través de los cuales las pequeñas moléculas hidrofílicas de los antibióticos beta-lactámicos alcanzarían las PBP de la membrana citoplasmática subyacente. Cualquiera que sea el mecanismo de penetración, lo cierto es que la concentración de determinados antibióticos beta-lactámicos requerida para inhibir el crecimiento de las bacterias gram-negativas son superiores a las necesarias para inactivar PBP esenciales en sistemas acelulares, incluso cuando no hay ninguna destrucción enzimática importante del antibiótico por beta-lactamasas bacterianas. No se requiere, empero, una concentración



ESTRUCTURA DE LA PARED CELULAR de bacterias gram-positivas (*izquierda*) y gram-negativas (*derecha*). La diferente estructura permite explicar en parte el hecho de que la penicilina *G* sea activa frente a las bacterias gram-positivas y no frente a las gram-negativas. En ambos casos, la membrana citoplasmática está cubierta por una capa de peptidoglicano, pero las bacterias gram-negativas tienen una membrana externa y pueden también almacenar enzimas con actividad beta-lactamasa en el espacio periplásmico, entre la membrana citoplasmática y la capa de peptidoglicanos. En la membrana citoplasmática hay proteínas que pueden unirse a la penicilina. Algunas, enzimas implicados en la síntesis de la pared celular, son sensibles al antibiótico.



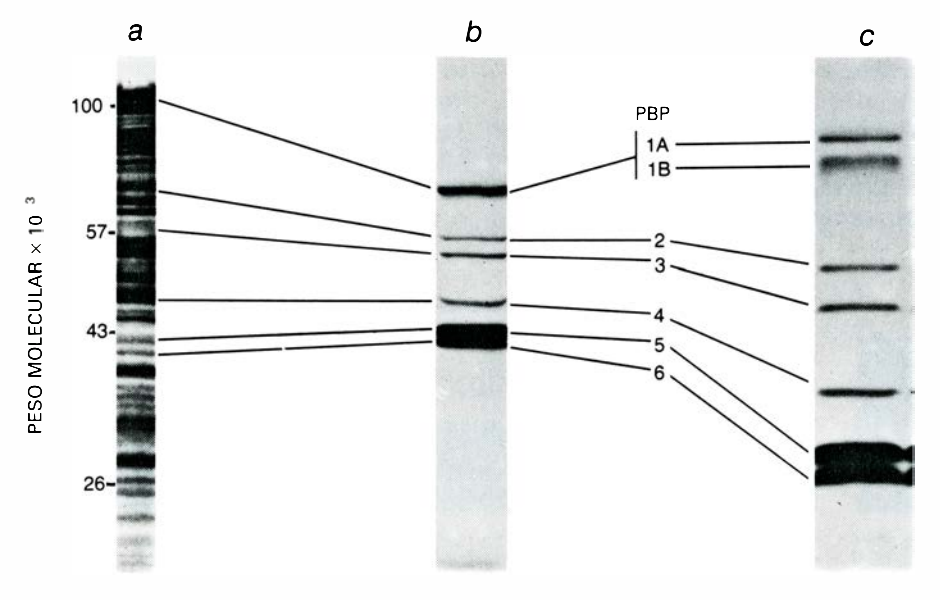
ESTRUCTURA DE LOS PEPTIDOGLICANOS de la pared bacteriana. Consiste en unidades alternantes de dos aminoazúcares, la N-acetilglucosamina (NAG) y el ácido N-acetilmurámico (NAM). Las unidades de NAM están unidas a cadenas peptídicas, muchas de las cuales se encuentran a su vez entrecruzadas. En la bacteria *Staphylococcus aureus*, el entrecruzamiento se consigue por la inserción de una glicina terminal en el enlace que une dos unidades de D-alanina de otra cadena, formando un nuevo enlace. Los enzimas transpeptidasa responsables de la inserción son atacados por las penicilinas y cefalosporinas cuando las bacterias se multiplican y forman nueva pared celular. (La ilustración es de Ilil Arbel.)

superior de antibiótico para inhibir el crecimiento de un mutante de *E. coli* cuya permeabilidad está alterada, y en el que hay libre acceso a la membrana.

Al parecer, la barrera que supone la existencia de la pared celular es, en parte, responsable de la relativa insensibilidad frente a muchos antibióticos beta-lactámicos que presenta la *Pseudomonas aeruginosa*, un organismo que produce infecciones graves en tejidos lesionados, y de la insensibilidad de muchas otras bacterias gram-negativas frente a penicilinas tales como la meticilina. La capacidad de las bacterias gram-negativas y gram-positivas de producir beta-lactamasas que pueden abrir el anillo beta-lactámico constituye otro factor primordial de la resistencia bacteriana y tiene una indudable importancia clínica. Este hecho se apreció por primera vez al observarse cepas de estafilococos productores de penicilinasa que seguían multiplicándose mientras las cepas sensibles se destruían. Estudios ulteriores han revelado que el enzima destructor de la penicilina lo segregan, en parte, las células bacterianas y, en parte, permanece unido a ellas. El enzima es inducido comúnmente por los antibióticos beta-lactámicos. Las poblaciones densas de estafilococos que producen grandes cantidades de beta-lactamasa destruyen rápidamente múltiples penicilinas, aun cuando las células individuales, que sintetizan sólo pequeñas cantidades, sigan siendo sensibles al antibiótico.

Muchas bacterias gram-negativas producen también una beta-lactamasa; en estos organismos el enzima es a menudo constitutivo, es decir, se sintetiza en grandes cantidades aunque no se añada un inductor. Se encuentra normalmente unido a las células, por más que suela vérsele en el espacio periplásmico, entre la membrana citoplasmática y la pared celular, desde donde se libera cuando se lesiona la pared. Ensayando la actividad de distintas beta-lactamasas procedentes de diversas bacterias frente a series de penicilinas y cefalosporinas, se observó que muchos de estos enzimas diferían en su capacidad para hidrolizar los diferentes antibióticos beta-lactámicos.

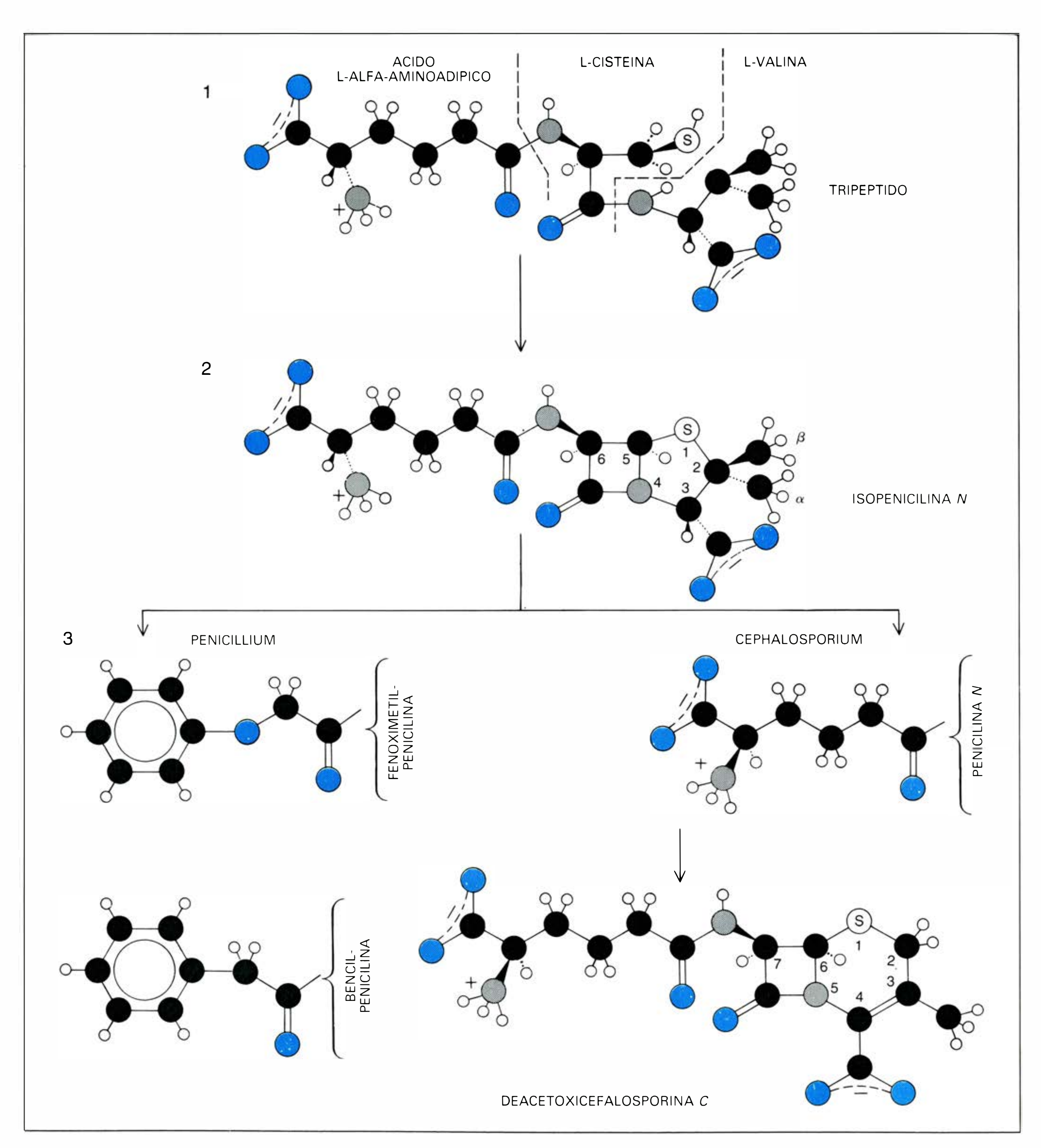
La multiplicidad de beta-lactamasas ha venido planteando serios problemas a los médicos. Se ha observado que los genes que codifican las beta-lactamasas pueden transferirse de una especie bacteriana a otra. Estos genes no son esenciales para la vida de las bacterias y, a menudo, forman parte de las unidades genéticas que experimentan autorreplicación conocidas como plásmidos. Richard P. Novick, del Instituto de Investigación de la Salud Pública de la ciudad de Nueva York, y Mark H. Richmond, de la Universidad de Bristol, demostraron que la producción de penicilinasa por estafilococos estaba mediada por plásmidos, y que podía transferirse de un estafilococo a otro incorporándola a la cabeza de un virus que infectase la célula. La transferencia a través de este mecanismo, conocido como transducción, puede tener poca importancia clínica. En las bacterias gram-negativas, sin embargo, los plásmidos pueden transferirse de una célula a otra de la misma especie, o de una especie diferente, por el proceso conocido como conjugación, en el que una célula dadora masculina se adhiere brevemente a una célula receptora femenina. Ade-



PROTEINAS CON AFINIDAD POR LA PENICILINA en las membranas citoplasmáticas de *E. coli*. Las proteínas fueron detectadas por Brian G. Spratt, de la Universidad de Sussex, mediante técnicas electroforéticas. La penicilina marcada con un isótopo radiactivo, el carbono 14, se unía a las membranas. Las proteínas se fraccionaron por electroforesis en un gel a través del cual migraban con diferentes velocidades, dependientes de su peso molecular. Las fotografías muestran el gel después de la tinción de las proteínas (*a*), las proteínas detectadas por autorradiografía (*b*) y la separación de las proteínas 1A y 1B con capacidad de unión de la penicilina por procedimientos electroforéticos de elevada resolución (*c*).

más, los genes que codifican la beta-lactamasa pueden encontrarse en fragmentos especializados del ADN, denominados transposones, susceptibles de translocarse de un plásmido a otro.

La capacidad de una célula bacteriana, o de un plásmido, de infectar a otra de su misma clase con genes que codifiquen una beta-lactamasa proporciona un mecanismo importante, en potencia, para la extensión de la resistencia. La resistencia no sólo puede transferirse de una bacteria patógena a otra, sino también de un organismo no patógeno a otro patógeno. En algunos casos, no obstante, la resistencia aparece lentamente. Durante muchos años, los gonococos permanecieron altamente sensibles a la bencilpenicilina, pero se ha identificado, en pacientes con gonorrea, una cepa resistente que poseía un



BIOSINTESIS de penicilinas y cefalosporinas a partir de los hongos *Penicillium* y *Cephalosporium*. La síntesis comienza (*1*) con la formación de un tripéptido, que consta, leyendo de izquierda a derecha, de tres unidades separadas por líneas interrumpidas: el ácido L-alfa-aminoadípico, la L-cisteína y la L-valina, respectivamente. El tripéptido se transforma en isopenicilina *N* (*2*). En las penicilinas, un enzima con actividad transferasa cataliza la sustitución de la cadena L-alfa-aminoadipil de la isopenicilina *N* (*3*), mientras que, en las cefalosporinas, no hay, aparentemente, transferasa, puesto que no se forman compuestos con una nueva cadena lateral. En lugar de ello, ocurren dos reacciones secuenciales. Primero, el centro asimétrico de la cadena lateral L-alfa-aminoadipil se invierte y da lugar a su imagen especular tridimensional, formándose la penicilina *N*. Después, la penicilina *N* se transforma en deacetoxicefalosporina por una expansión del anillo, que supone la ruptura del enlace existente entre el átomo de azufre y el segundo carbono, y la unión de un grupo beta-metilo al átomo de azufre a través del carbono del grupo beta-metilo. Se desconoce el mecanismo de las distintas reacciones.

plásmido con el gen de la beta-lactamasa similar al descrito en otras bacterias gram-negativas.

¿Qué resultados se han obtenido hasta el presente en la investigación de los antibióticos beta-lactámicos con vistas a resolver los problemas clínicos? Un avance significativo partió del descubrimiento realizado en Oxford, en 1953, de que un hongo del género *Cephalosporium*, que había sido aislado en 1945 por G. Brotzu, en Cerdeña, producía una nueva penicilina con una cadena lateral derivada del ácido D-alfa-aminoadípico. G. G. F. Newton y el autor descubrieron que las preparaciones crudas de esta penicilina contenían pequeñas cantidades de un segundo antibiótico con la misma cadena lateral, aunque con una característica estructural no advertida hasta entonces: el anillo beta-lactámico estaba unido a un anillo de 6 átomos, en vez de 5. El nuevo antibiótico, denominado cefalosporina *C*, tenía interés clínico, puesto que poseía, junto a la característica de baja toxicidad propia de las penicilinas, resistencia a la hidrólisis por determinadas beta-lactamasas.

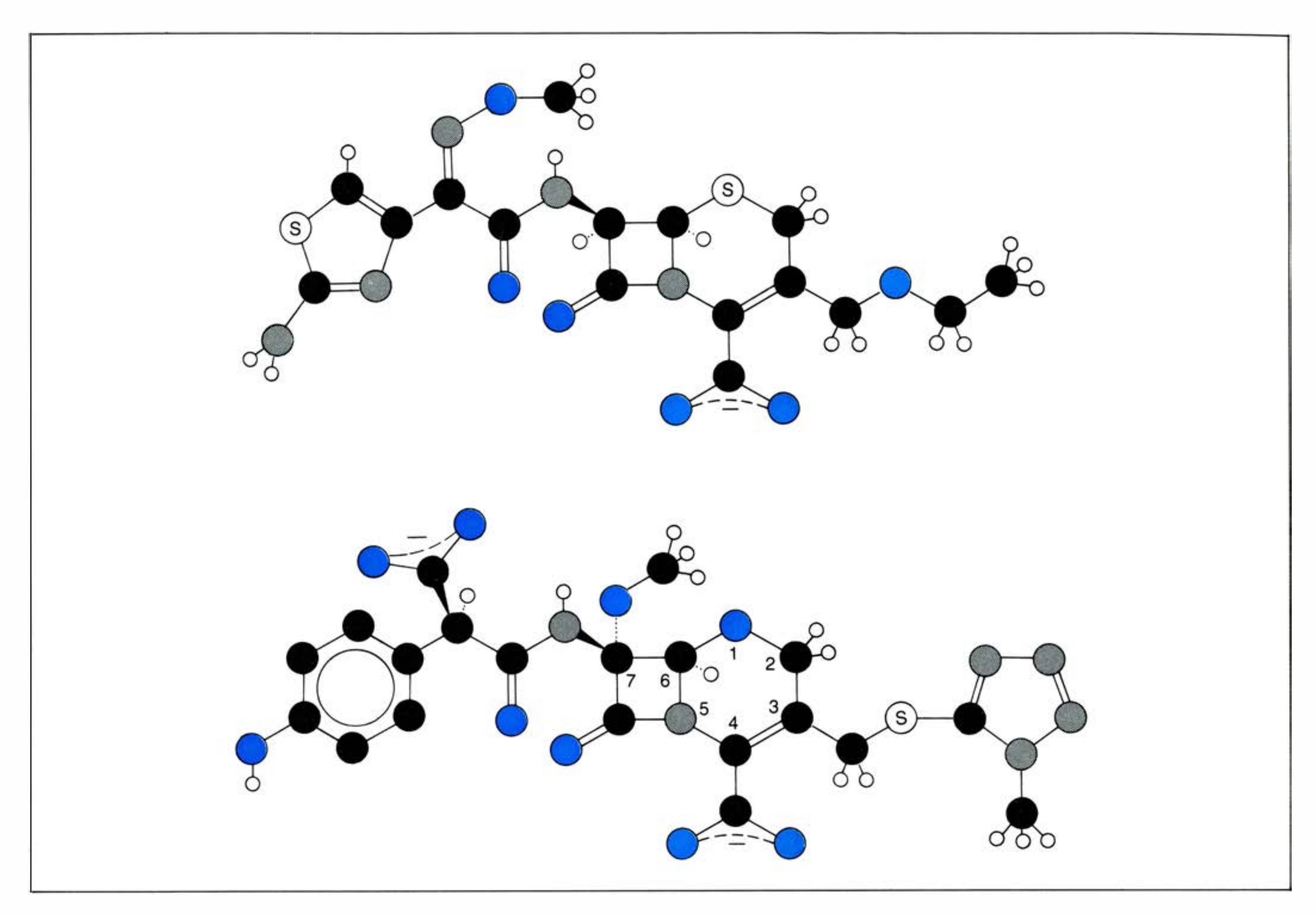
Penicillium y *Cephalosporium* requieren, para su cultivo en laboratorio o en una planta industrial, distintas condiciones de fermentación. Ambos hongos, sin embargo, siguen vías de biosíntesis similares para producir antibióticos beta-lactámicos. Los estudios llevados a cabo en preparaciones acelulares han demostrado que una cadena de tres aminoácidos forma un anillo, dando lugar a la isopenicilina *N*, una penicilina con una cadena lateral constituida por un residuo L-alfa-aminoadipil. En el *Penicillium chrysogenum*, una aciltransferasa cataliza posteriormente la sustitución de la citada cadena lateral por un derivado formado a partir de un precursor, ácido fenilacético, por ejemplo. Las especies de *Cephalosporium* carecen de la transferasa y no se forma ningún compuesto con una nueva cadena lateral. En su lugar, ocurren dos reacciones secuenciales. El centro asimétrico de la cadena lateral se invierte y da lugar a su imagen especular tridimensional. Parte de la penicilina *N* resultante, con una cadena lateral D-alfa-aminoadipil, se segrega, pero parte se transforma en deacetoxicefalosporina *C*, que origina, ulteriormente, cefalosporina *C* y otras cefalosporinas, según han demostrado los recientes experimentos llevados a cabo por Arnold L. Demain y sus colaboradores, del MIT.

Por simple fermentación sólo se han podido obtener ciertos tipos de penicilina y cefalosporinas con la cadena lateral D-alfa-aminoadipil. Por métodos semisintéticos, sin embargo, se han preparado y estudiado miles de compuestos de la familia de los antibióticos beta-lactámicos. La semisíntesis resultó posible con el descubrimiento de métodos para la producción del núcleo de la molécula de penicilina, el ácido 6-aminopenicilánico (6-APA) y del de la cefalosporina, el ácido 7-alfa-aminocefalosporánico, que carecen del residuo acil lateral. Partiendo de informes japoneses sobre la producción de 6-APA por el *Penicillium chrysogenum*, John Sheehan logró su síntesis; de forma independiente se descubrió en los laboratorios de investigación Beecham, de Gran Bretaña, donde se obtuvo a partir de fermentaciones a las que no se había añadido precursor de la cadena lateral. Posteriormente se produjo en mayores cantidades eliminando la cadena lateral de la bencilpenicilina con acilasas bacterianas.

No se encontraron enzimas capaces de eliminar la cadena lateral D-alfa-aminoadipil de la cefalosporina *C*. Pese a ello, la 7-ACA se produjo en pequeñas cantidades en Oxford mediante hidrólisis ácida suave de la cefalosporina *C* y, posteriormente, siguiendo un procedimiento químico mucho más eficiente utilizado en los laboratorios de investigación Eli Lilly. Estos logros abrieron las puertas a la obtención de penicilinas semisintéticas y cefalosporinas enlazando, por métodos químicos, distintas cadenas laterales al 6-APA y 7-ACA. La molécula de cefalosporina pudo modificarse más adelante sin alterar su anillo esencial mediante procedimientos químicos que intercambiaban el grupo acetoxi (-$OCOCH_3$), unido al átomo 3' que se proyecta del anillo de 6 átomos, por otros grupos o por un átomo de hidrógeno.

En la investigación de las cefalosporinas merecen destacarse otros dos hallazgos. Uno de ellos, descubierto en los laboratorios Lilly, consistía en una ingeniosa vía química para producir el anillo de la cefalosporina a partir del anillo de la penicilina, lo que permitió obtener ciertas cefalosporinas utilizando penicilinas como sustrato. Por otro lado, los laboratorios Lilly y los laboratorios Merck Sharp & Dohme descubrieron que la cefalosporina obtenida a partir de *Streptomyces* poseía un grupo metoxil (-OCH_3) en lugar del átomo de hidrógeno diagonalmente opuesto al nitrógeno del anillo beta-lactámico. Posteriormente, se han desarrollado métodos para introducir el

NUEVOS ANTIBIOTICOS BETA-LACTAMICOS descubiertos a partir de 1976, entre los que se incluyen la tienamicina (*a*), el ácido clavulánico (*b*) y los penems, uno de los cuales se muestra en la figura (*c*). La tienamicina y el ácido clavulánico derivan de *Streptomyces*. Los antibióticos poseen un anillo beta-lactámico, pero no el característico de la penicilina o de la cefalosporina. Los penems fueron sintetizados por R. B. Woodward, de la Universidad de Harvard, y se caracterizan por poseer el doble enlace del anillo que contiene azufre de las cefalosporinas en el anillo correspondiente de las penicilinas.



TERCERA GENERACION DE CEFALOSPORINAS, entre las que se encuentran la cefotaxima (*arriba*) y el moxalactam (*abajo*), una 1-oxacefalosporina derivada por semisíntesis de la penicilina. Una característica de la 1-oxacefalosporina es que el oxígeno sustituye al azufre en el anillo de seis átomos.

grupo metoxil en otras moléculas de cefalosporina.

Se han obtenido miles de penicilinas y cefalosporinas por uno u otro de los métodos citados; algunas revistieron una importancia clínica innegable. La primera nueva penicilina de interés médico, denominada meticilina, se produjo en los laboratorios Beecham; resultó activa frente a los estafilococos resistentes a la bencilpenicilina, debido a que su voluminosa cadena lateral obstruía la interacción con la penicilinasa estafilocócica. Otros productos fueron la ampicilina, que mostró una mayor actividad frente a un cierto número de bacterias gram-negativas, y la carbenicilina, útil por su actividad frente a la bacteria *Pseudomonas aeruginosa*.

Las dos primeras cefalosporinas semisintéticas, la cefalotina de Lilly y la cefaloridina, de los laboratorios británicos Glaxo, resultaron eficaces frente a los estafilococos resistentes a la penicilina; pero en este caso era la naturaleza del anillo, y no la de la cadena lateral, lo que protegía los antibióticos de la hidrólisis rápida por el enzima estafilocócico. Las dos cefalosporinas mostraron también actividad frente a varias bacterias gram-negativas, sin provocar reacciones en la mayoría de pacientes con hipersensibilidad a la penicilina. Aunque el anillo de las cefalosporinas es bastante estable en el medio ácido gástrico, las primeras cefalosporinas no se absorbían al administrarse por vía oral. Y sin que nadie lo esperase, otra cefalosporina semisintética, la cefalexina, resultó efectiva por vía oral.

A pesar del notable valor terapéutico que tenían las primeras penicilinas y cefalosporinas semisintéticas, algunas bacterias patógenas gram-negativas eran insensibles a las mismas. Nuevos cambios en los grupos enlazados al anillo han dado lugar a la segunda y tercera generación de cefalosporinas. Los primeros tres productos se denominaron cefamandol, cefuroxima y cefoxitina; les siguieron la cefotaxima y cefoperazona. Por semisíntesis a partir de la penicilina se obtuvo una molécula altamente activa, la 1-oxacefalosporina, con oxígeno en lugar de azufre en el anillo de 6 átomos. Muchos de estos compuestos han supuesto un avance considerable en la quimioterapia bacteriana, por su espectro de actividad más amplio que el de los antibióticos beta-lactámicos más antiguos, y su mayor resistencia a la hidrólisis desencadenada por múltiples beta-lactamasas.

Un nuevo capítulo en la historia de los antibióticos beta-lactámicos se abrió en 1976 con el descubrimiento, en varios laboratorios farmacéuticos, de que ciertos *Streptomyces* producían sustancias que poseían un anillo beta-lactámico, pero no el anillo de la penicilina ni el de la cefalosporina. Algunas de las nuevas sustancias se detectaron en ensayos de actividad antibacteriana y otras en la búsqueda de sustancias con capacidad inhibidora de la acción de las beta-lactamasas. Una de ellas, la tienamicina, que muestra un amplio espectro antibacteriano y alta resistencia a las beta-lactamasas, puede resultar de interés clínico. Otra, el ácido clavulánico, constituye un potente inactivador de las beta-lactamasas. Al mismo tiempo, Robert Woodward describió la síntesis y semisíntesis de anillos beta-lactámicos denominados penems, en los que el doble enlace del anillo que contiene azufre de las cefalosporinas se ha colocado en el anillo correspondiente a las penicilinas. Muy recientemente se han aislado nuevos antibióticos beta-lactámicos procedentes de medios de cultivo de distintas especies bacterianas.

¿Qué puede predecirse a propósito de los nuevos antibióticos beta-lactámicos y su relación con la medicina clínica? A veces no ha resultado difícil pronosticar el alcance de un producto. Una cadena lateral capaz de atraer un electrón aumentaba la estabilidad de la penicilina en presencia de ácido, de modo que el medicamento resultaba eficaz al administrarlo por vía oral. Por analogía con las conocidas diferencias entre la penicilina *N* y la bencilpenicilina, parecía probable que el intercambio de la cadena lateral alfa-aminoadipil de las cefalosporinas naturales por cadenas laterales menos polares (cadenas desprovistas de cargas positivas y negativas) daría lugar a un gran incremento de la actividad intrínseca frente a algunas bacterias y al mantenimiento de su estabilidad frente a la penicilinasa estafilocócica.

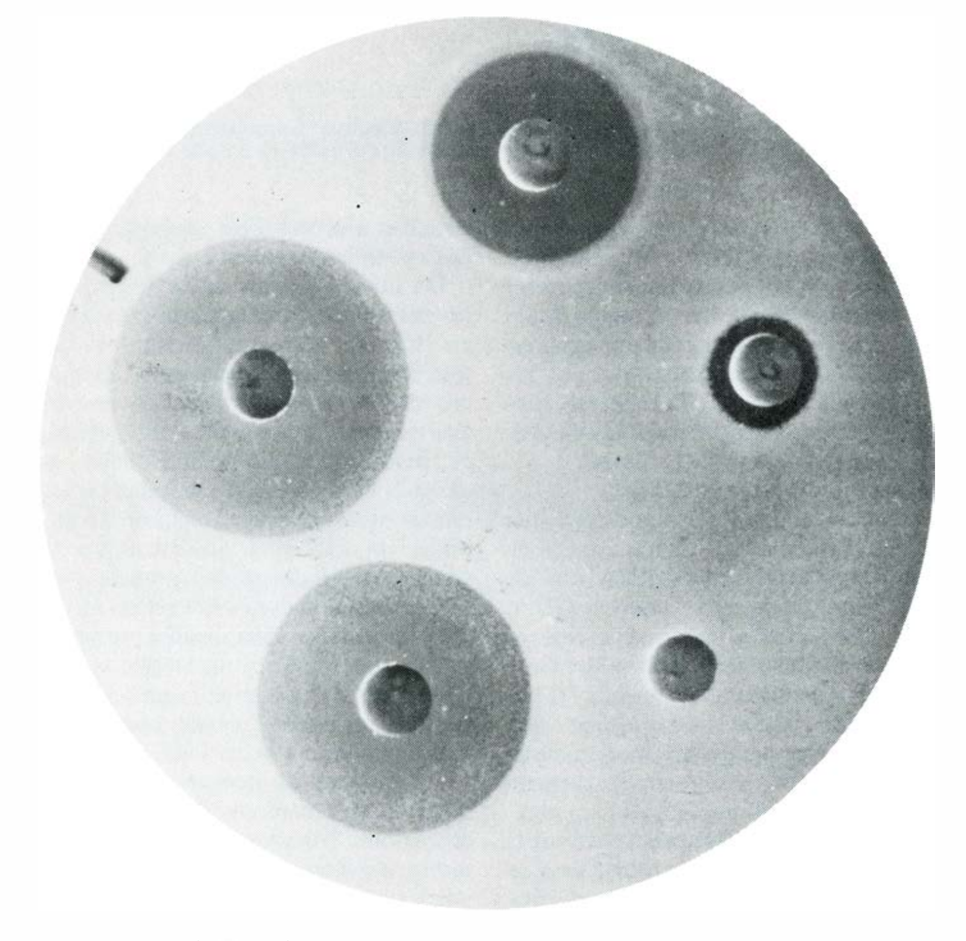
Así fue. La observación de que el grupo O-acetil del carbono 3' de las cefalosporinas, al que se ha hecho referencia anteriormente, podía eliminarse en el organismo por una acetil esterasa, dando lugar a compuestos menos activos, sugirió la conveniencia de sustituir este grupo por otros más estables. No han podido determinarse las características estructurales de las cefalosporinas requeridas para su absorción intestinal; ni tampoco se conocen las propiedades requeridas para que un antibiótico beta-lactámico sea un inactivador potente de enzimas de la membrana bacteriana y, al mismo tiempo, resista la hidrólisis que provocan muchas beta-lactamasas. Por todo ello, si bien se han obtenido miles de compuestos, sólo unos pocos han resultado de utilidad clínica. Algunos de ellos se conocieron gracias a observaciones casuales; otros, tras exhaustivos ensayos sistemáticos.

Para desvelar el mecanismo de acción de los antibióticos beta-lactámicos resulta esencial el conocimiento de la estructura tridimensional de los sitios activos de los distintos enzimas con los que reaccionan estos antibióticos. Se están identificando ya pormenores significativos acerca de la naturaleza y función de tales sitios.

Por cromatografía de afinidad pueden purificarse proteínas con capacidad de unión a la penicilina y con actividad D-carboxipeptidasa. En este ensayo se utilizan penicilinas o cefalosporinas adecuadas, que, previamente, se han unido a una resina. Las PBP quedan retenidas específicamente en la columna y pueden eluirse con hidroxilamina, que rompe la unión covalente formada por la reacción de la PBP con el anillo beta-lactámico. Siguiendo este procedimiento, se han determinado las secuencias de aminoácidos de fragmentos de dos PBP notables y se ha identificado el sitio de unión de la penicilina. La purificación de otras PBP menores que poseen, presumiblemente, actividad transpeptidasa y las inhiben los antibióticos beta-lactámicos, ha resultado hasta ahora difícil; se han dado, empero, pasos importantes mediante la inserción en *E. coli* de ADN que determina la producción de múltiples copias del gen que codifica para una PBP. Se han realizado insertando los genes en un plásmido o en un virus que producen múltiples copias del gen cuando infectan la célula bacteriana y se multiplican.

Se ha logrado ya la purificación de diversas beta-lactamasas. Gracias a los trabajos realizados por Richard P. A. Ambler, de la Universidad de Edimburgo, se conocen las secuencias de aminoácidos de enzimas procedentes de *Staphylococcus aureus, Bacillus licheniformis* y *B. cereus*. La secuencia de una beta-lactamasa de *E. coli* codificada por un plásmido ha sido deducida por J. G. Sutcliffe, de Harvard, a partir de la secuencia de nucleótidos y el correspondiente gen. Salvo una excepción (un enzima de *B. cereus* que requiere zinc), las secuencias muestran una considerable homología, es decir, muestran mayor número de similaridades de las que pueden atribuirse al azar. Se ha observado también estrecha homología entre beta-lactamasas y dos PBP con actividad D-carboxipeptidasa. Es posible, pues, que un grupo de beta-lactamasas y un grupo de enzimas sensibles a la penicilina de las membranas bacterianas tengan un origen evolutivo común.

Estas dos familias de enzimas muestran también similaridades en su modo de acción. Utilizando compuestos de bencilpenicilina que permanecen unidos a sus sitios activos, se ha demostrado que el grupo hidroxilo de la serina ataca el anillo beta-lactámico. El ataque conduce a la apertura del anillo, con formación de un acilenzima, un éster de la serina. En presencia de una beta-lactamasa, el acil-enzima de muchos antibióticos beta-lactámicos se hi-



EFECTO DE LOS PENEMS, puesto de manifiesto en esta fotografía; se aprecia la acción inhibidora de varios antibióticos frente al *Staphylococcus aureus*, bacteria resistente a la penicilina, cultivado en agar. Cada uno de los cinco círculos oscuros representa un área donde las bacterias han sido destruidas por un determinado antibiótico; cuanto mayor es el círculo, más efectivo es el antibiótico. En el sentido de las agujas del reloj, desde la parte superior, los antibióticos son: la cefaloxina; la penicilina *V*; un penem inactivo (en el que el azufre se proyecta por debajo del plano del carbono beta-lactámico en lugar de hacia arriba, como en las penicilinas y cefalosporinas, originando una forma tridimensional anormal); una mezcla de penems con configuraciones naturales y anómalas y un penem activo de configuración normal.

droliza rápidamente, regenerándose el enzima y dando lugar a un peniciloato, o un compuesto análogo, desprovisto de actividad antibacteriana. En presencia de una carboxipeptidasa de la membrana bacteriana, la hidrólisis, del acil-enzima del antibiótico beta-lactámico es muy lenta, lo que explica la acumulación del enzima inactivo.

En una D-carboxipeptidasa, la serina que reacciona con la penicilina es la que reacciona con un sustrato del enzima que contiene D-alanil-D-alanina. Esta observación apoya la hipótesis propuesta en 1965 por Jack Strominger y Donald J. Tipper, de la Universidad de Wisconsin en Madison. Para explicar por qué la penicilina inhibe el entrecruzamiento en la pared celular, postularon que la posición en el espacio de los átomos de la penicilina mimetizaba la posición de los átomos del extremo D-alanil-D-alanina de las cadenas laterales de los peptidoglicanos. Se desconoce en qué medida puede aplicarse esta hipótesis a la hora de justificar las diferentes capacidades de los antibióticos beta-lactámicos para inhibir el funcionamiento de PBP esenciales de la membrana.

Entre los nuevos compuestos beta-lactamasas se encuentran algunos cuya interacción con el enzima produce su na a la hidrólisis producida por las beta-lactamsas se encuentran algunos cuya interacción con el enzima produce su inactivación temporal o permanente. Este fenómeno lo están estudiando Jeremy Knowles, de Harvard, Stephen G. Waley, de Oxford, y otros. Si la velocidad de hidrólisis del acil-enzima intermediario es más lenta que su formación, se producirá un acúmulo del intermediario y el sitio activo no estará disponible para un sustrato fácilmente hidrolizable, tal como la bencil-penicilina. Cuando el sustrato que se hidroliza lentamente se ha agotado, y no puede formarse nuevo acil-enzima, el enzima activo puede regenerarse. En algunos casos, se produce un cambio estructural en el grupo acilo de la molécula del acil-enzima que bloquea por completo la liberación de tal grupo. Cuando esto sucede, puede decirse que el enzima se ha suicidado.

La idea de que un antibiótico beta-lactámico pueda proteger a otro de la destrucción por una beta-lactamasa arranca de observaciones realizadas en 1956, cuando la hidrólisis de la bencil-penicilina por una beta-lactamasa resultaba competitivamente inhibida por la cefalosporina *C*. No se ha generalizado la administración de mezclas de dos compuestos que actúan de forma sinérgica por esta razón. No obstante, después del descubrimiento de que el ácido clavulánico y otras sustancias no sólo inhiben las beta-lactamasas competitivamente, sino que las inactivan irreversiblemente, habrán de explorarse de nuevo las posibilidades de tal sinergismo.

Los sorprendentes avances en el campo de los antibióticos beta-lactámicos a partir del descubrimiento del valor terapéutico de la penicilina, y la intensa investigación que se realiza en este campo, se han visto estimulados en gran medida por el extraordinario valor terapéutico de estos antibióticos desprovistos de toxicidad. Se ha avanzado mucho en la solución de los problemas que plantean las bacterias resistentes a la penicilina, si bien los cambiantes patrones de resistencia y la extraordinaria mutabilidad de los microorganismos determinan que cualquier solución sea meramente provisional.

Hace 30 años, la molécula de penicilina parecía englobar las propiedades de un anillo único. Actualmente, el único factor común en la estructura de los antibióticos beta-lactámicos es el anillo beta-lactámico. El átomo de nitrógeno del anillo suele estar unido a un átomo de carbono con un grupo carboxilo (-COOH), un grupo ionizable emparentado, de forma similar, o una lactona (un éster cíclico de un ácido hidroxílico). Pero en los nuevos antibióticos beta-lactámicos, el nitrógeno forma parte de un grupo sulfamato ($-NSO_3^-$). Para que la molécula muestre una elevada actividad intrínseca frente a bacterias por formación de un acil-enzima, parece que el anillo beta-lactámico habría de poseer una elevada actividad química. Ello se logra en las penicilinas y cefalosporinas por fusión del anillo de 4 átomos beta-lactámico en un segundo anillo, lo que da lugar a un sistema en el que la configuración normal resulta forzada.

Si el anillo beta-lactámico es demasiado reactivo, la molécula carecerá de la estabilidad necesaria para su aprovechamiento clínico, y sus efectos biológicos dejarán de ser altamente específicos. Además, si los grupos unidos al anillo no ocupan posiciones adecuadas en el espacio, la molécula no encajará en los sitios activos de las proteínas que se engarzan a la penicilina en la membrana celular, de modo que pueda formarse un acil-enzima relativamente estable. Cabe esperar que en los próximos diez años podamos especificar con detalle las características estructurales que deben reunir los antibióticos beta-lactámicos potencialmente útiles.

Ciencia y sociedad

Microgravedad y biología

Desde el comienzo de la evolución biológica los seres vivos se desarrollan, existen y mueren bajo la influencia permanente de fuerzas cuyo origen se encuentra en las propiedades físicas y geológicas de nuestro planeta. A modo de ejemplo podemos mencionar algunas de las más importantes: la radiación electromagnética, solar y cósmica, el campo magnético terrestre, los campos eléctricos atmosféricos, el campo gravitatorio terrestre y la presión atmosférica. Para el ser vivo, que en general habita en un lugar relativamente constante, algunas de estas fuerzas están siempre presentes (la fuerza de la gravedad), otras cambian relativamente poco (campo magnético, presión atmósferica), mientras que otras cambian muy marcadamente, por ejemplo, la radiación electromagnética solar. Estos cambios en muchos casos no son aleatorios, sino que siguen ciclos precisos, más o menos cortos.

La pregunta que surge es si los seres vivos han utilizado y utilizan hoy en día alguna o todas estas fuerzas para dirigir y regular su comportamiento, desarrollo y alguno o muchos de los procesos fisiológicos cuyo conjunto produce esa maravillosa diversidad y constancia que encontramos en la biosfera. Aunque en el caso de alguna de estas fuerzas geofísicas existen suficientes pruebas experimentales para contestar afirmativamente a esta pregunta, en el caso de otras sabemos muy poco sobre sus posibles influencias, incluso si ni siquiera actúan realmente sobre nosotros. A priori, se puede adelantar que sería sorprendente que no existiera tal influencia, pues una de las características fundamentales de los seres vivos ha sido y es su capacidad de utilizar y explotar todos los elementos, procesos y fuerzas físico-químicas del medio ambiente donde han evolucionado y siguen evolucionando hoy en día. En el caso de algunos de estos factores, como la gravitación terrestre, su presencia constante dificulta la investigación de sus posibles efectos. Aunque existen algunas técnicas de simulación en el laboratorio, y sobre todo la posibilidad de someter a los seres vivos a fuerzas de gravedad aumentadas, las limitaciones de estas técnicas experimentales no hacen más que reforzar la necesidad de realizar pruebas en condiciones de ausencia total o prácticamente total de gravedad, si queremos llegar a saber cómo y sobre qué procesos biológicos actúa.

El desarrollo de la exploración espacial y los proyectos de establecimiento de laboratorios y estaciones espaciales permanentes o cuasipermanentes permiten abordar estos problemas de forma controlada. Y así, la agencia espacial europea (ESA) está empezando a desarrollar un programa científico específico, el programa de microgravedad. Este programa tiene dos vertientes: por un lado incluye problemas físicos y tecnológicos, como es el estudio del comportamiento de fluidos y otros procesos que afectan a los sistemas metalúrgicos como el crecimiento de cristales, etcétera, y, por otro lado, la investigación del papel que la gravedad desempeña en cualquier aspecto relacionado con los seres vivos.

Aunque a nadie se le oculta que el desarrollo de la exploración espacial ha sido principalmente por motivos de prestigio e incluso de poderío militar de las naciones desarrolladas, este esfuerzo ha tenido consecuencias tecnológicas importantes en muchas áreas. También ha sido aprovechado por la investigación espacial que podríamos llamar tradicional, que afecta sobre todo a las ciencias astronómicas y astrofísicas, así como a la meteorología. Por otra parte, los temas biológicos han estado presentes desde un principio en los vuelos espaciales tanto soviéticos como norteamericanos. Ya en 1948 la armada de los Estados Unidos hizo volar a Alberto, un mono, en un cohete V-2. Sin embargo, la preocupación principal de estos experimentos era detectar, y eventualmente prevenir o curar, cualquier efecto deletéreo que las condiciones del vuelo espacial, incluyendo la falta de gravedad, pudieran tener sobre la persona humana. Como es obvio, esto tenía una razón inmediata de aplicación, y sólo de modo accesorio se realizaron experimentos que intentaban aclarar los mecanismos básicos del posible efecto de la gravedad o de otros parámetros alterados durante el vuelo espacial sobre los seres vivos. Se encontraron algunos fenómenos que no podían explicarse en términos de la experiencia en tierra, pero por un número de razones, incluyendo la forma elemental y poco controlada de los experimentos, las conclusiones que se pueden sacar de los mismos en el mejor de los casos sólo son provisionales.

La agencia espacial europea es un organismo internacional que nació en 1975 de la fusión de dos organismos europeos que existían desde 1964, el ESRO, responsable del desarrollo de satélites europeos, y el ELDO, responsable del desarrollo de vehículos de lanzamiento. Son miembros plenos: Alemania Federal, Bélgica, Dinamarca, España, Francia, Gran Bretaña, Holanda, Irlanda, Italia, Suecia y Suiza. Son miembros asociados: Austria y Noruega; Canadá ha pedido figurar como miembro en estrecha asociación. El objetivo de la ESA es la cooperación de los estados europeos en el desarrollo, con propósitos exclusivamente pacíficos, de la investigación y tecnología espaciales y sus aplicaciones. Esta organización cuenta con un presupuesto de unos 850 millones de dólares para 1981. España participa desde su inicio en esta organización multinacional europea contribuyendo para este año con un 2,42 por ciento de este presupuesto, cantidad respetable si se considera el bastante menor esfuerzo que nuestro país viene tradicionalmente dedicando a gastos de investigación y desarrollo.

Dentro de la filosofía general de la agencia espacial europea está la construcción del Spacelab o laboratorio espacial. El Spacelab es la contribución europea al sistema de transporte espacial avanzado de la NASA (la agencia espacial norteamericana). Es un laboratorio espacial tripulado, donde por primera vez científicos e ingenieros, en vez de astronautas, trabajarán en el espacio llevando a cabo experimentos en una amplia gama de disciplinas.

Cuando en 1978 la ESA hizo una convocatoria preliminar de proyectos de investigación para el Spacelab, la aparición de un número importante de proyectos en las ciencias biológicas llevó a plantear el desarrollo de una unidad versátil y reutilizable para experimentos biológicos, capaz de aceptar múltiples experimentos simultáneamente. Esta unidad lleva el nombre de Biorack. La decisión final para llevar adelante el desarrollo de esta unidad, es decir, del Biorack, se condicionó a que existiera un suficiente interés por parte de la comunidad científica europea en este tipo de experimentos. Para ello se hizo una nueva convocatoria, ahora específicamente dirigida a las ciencias biológicas, durante los meses de la primavera de 1980. Los proyectos de investigación recibidos han seguido el tratamiento normal que suelen recibir en los países científicamente desarrollados, es decir, han sido evaluados críticamente por unos paneles independientes de expertos: uno en el área de biología celular y del desarrollo, otro

en biología vegetal y un tercero en el área de radiología. La ESA recibió 56 proyectos de los que los comités de expertos seleccionaron 26 que juzgaron contenían propuestas capaces de contribuir a la solución de cuestiones biológicas que requieren experimentación en condiciones no existentes en la Tierra. Los comités de expertos llegaron a la conclusión global de que la calidad y número era más que suficiente para seguir adelante con el desarrollo del Biorack. Expresaron su opinión de que la investigación espacial, y en particular la de microgravedad, puede permitir la exploración de nuevos aspectos de la biología. Esos comités han quedado impresionados por la variedad y alcance de los proyectos presentados, especialmente teniendo en cuenta el pequeño número de biólogos europeos con experiencia previa en la investigación espacial.

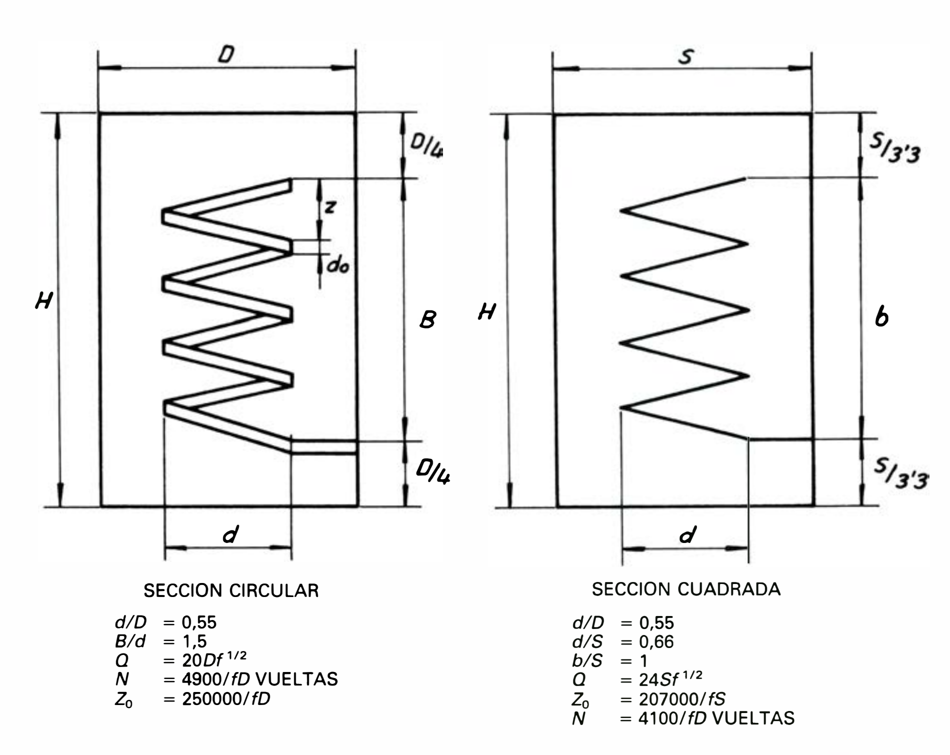
Nuestro grupo del Instituto de Enzimología y Patología Molecular del Consejo Superior de Investigaciones Científicas presentó un proyecto sobre la participación de la fuerza de la gravedad en la distribución topológica de los determinantes citoplasmáticos en el huevo de la mosca del vinagre, *Drosophila melanogaster*. El proyecto ha sido juzgado muy favorablemente por los comités de expertos y seleccionado por tanto como prioritario para uno de los próximos vuelos del Spacelab. Su contenido se inserta dentro de las líneas de investigación de nuestro grupo, dirigidas a aclarar uno de los procesos biológicos más importantes aún por entender, el del desarrollo embrionario temprano. Este problema, que está con nosotros desde siempre y que el hombre reconoció desde que empezó a hacer observaciones científicas, permanece aún por resolver a pesar de los enormes progresos de las ciencias biológicas en los últimos decenios. El misterio de cómo se generan con precisión y regularidad organismos desarrollados complejos muy diferentes, a partir de sistemas semejantes, aparentemente simples y uniformes, como son la mayoría de los huevos fecundados, resta aún por aclarar en sus aspectos más fundamentales. La embriología experimental clásica ha acumulado considerables pruebas sobre la existencia de unos presuntos determinantes citoplasmáticos cuya función estaría en señalizar y poner en marcha el desarrollo de las distintas partes del embrión. La naturaleza molecular de estos determinantes citoplasmáticos se sigue desconociendo, a pesar de haber sido buscados con bastante intensidad. Tampoco se sabe qué fuerzas conducen a su distribución no uniforme en el citoplasma del huevo. Desde finales del siglo pasado se ha venido sugiriendo que la gravedad podría ser una de las fuerzas que los ovocitos y embriones tempranos pueden utilizar con este fin. Las pruebas experimentales sobre ello siguen siendo conflictivas y en este sentido los experimentos propuestos a realizar en el Spacelab podrían contribuir decisivamente a su confirmación o su rechazo definitivo.

El concepto del Spacelab abre una nueva dimensión a la investigación espacial. Aquí se pueden programar verdaderos experimentos controlados que los especialistas de la tripulación pueden seguir durante el tiempo del vuelo realizando las manipulaciones y operaciones necesarias para su desarrollo. Por ejemplo, varios de los proyectos presentados, incluyendo el nuestro, insistían en la importancia de llevar en paralelo durante el vuelo un experimento control en condiciones de gravedad normal. Así se podrá reconocer si un efecto cualquiera observado se debe a la falta de gravedad o a cualquiera de las otras condiciones alteradas durante el vuelo espacial. Este control es tan importante que los comités de expertos han recomendado que se incluya obligatoriamente y ha llevado a los técnicos de la agencia a incluir dentro del diseño del Biorack una serie de centrífugas de baja velocidad donde simular las condiciones de gravedad normal. Pues bien, ningún experimento anterior, ni soviético ni norteamericano, ha incluido este tipo de control, lo que ilustra las limitaciones que la investigación espacial ha tenido en el pasado. (Roberto Marco.)

Filtros con respnadores helicoidales

El gran interés por los equipos de radiocomunicaciones móviles en las bandas de frecuencias de VHF y UHF nos movió, a un grupo del Centro de Investigación y Estudios de la Compañía Telefónica Nacional de España, a plantear un trabajo de investigación sobre el diseño de filtros de radiofrecuencia de alta selectividad, en dichas bandas, haciendo hincapié en aspectos de tipo industrial. El estudio se centró en minimizar el volumen de los filtros y en reducir el tiempo de ajuste durante el proceso de fabricación. Las técnicas convencionales para la realización de filtros de radiofrecuencia consisten en la utilización de resonadores en línea coaxial en cuarto de longitud de onda, cuya dimensión axial oscila entre 75 y 15 centímetros para las bandas de frecuencias de VHF y UHF, generalmente empleadas para los sistemas de radiocomunicaciones móviles terrestres. La reducción del volumen de dichos filtros se consigue introduciendo una carga capacitiva; esta técnica, sin embargo, dificulta la obtención de filtros de alta selectividad, aumenta las pérdidas por disipación y dificulta el ajuste durante el proceso de fabricación.

El desarrollo de la microelectrónica



*Resonador helicoidal de sección circular (**izquierda**) y cuadrada (**derecha**). Las dimensiones se dan en centímetros y la frecuencia en megahertz*

permite reducir drásticamente el peso y el volumen de la circuitería electrónica asociada a los equipos de telecomunicación, lo que se refleja en su aplicación a los equipos móviles. El mayor inconveniente para la miniaturización de estos equipos estriba en la introducción de filtros de alta selectividad, cuya necesidad se hace cada vez más imperiosa a medida que proliferan los equipos móviles y en razón de los requisitos que impone el correcto aprovechamiento del espectro. La búsqueda de procedimientos para reducir las dimensiones de los filtros de radiofrecuencia condujo a investigar sobre la posibilidad de utilizar resonadores en línea de transmisión helicoidal. La línea de transmisión helicoidal consiste en reemplazar el conductor interior de la línea coaxial por un conductor bobinado en hélice. Con esta estructura se consigue reducir la longitud de onda efectiva en el interior de la línea, lo que permite construir resonadores en cuarto de longitud de onda de menores dimensiones que los realizados con técnicas usuales.

En el resonador helicoidal se combinan tanto los parámetros concentrados como los distribuidos. La línea de transmisión es una estructura coaxial donde el conductor interior es una hélice y el conductor exterior puede tener una sección cualquiera, aunque se utiliza con preferencia la sección circular o cuadrada. Su ventaja con respecto a las líneas coaxiales convencionales radica en que la velocidad de propagación a lo largo de la línea de transmisión es inferior a la que existe en el aire, por ser una estructura periódica. Así, la longitud de onda es inferior a la correspondiente al modo TEM en una línea coaxial y, por tanto, la longitud física de un resonador en cuarto de longitud de onda en línea helicoidal disminuye con respecto a su equivalente realizado sobre línea de transmisión coaxial convencional.

En la figura izquierda de la ilustración de la página precedente se muestra el esquema de un resonador helicoidal de sección circular y las fórmulas de diseño; en la figura de la derecha se indican las correspondientes al resonador helicoidal de sección cuadrada. El ajuste de la frecuencia de resonancia puede realizarse mecánicamente, mediante un tornillo metálico en conexión con la tapa superior, que introduce una capacidad con el extremo libre de la hélice del resonador y desplaza la frecuencia de resonancia en torno al valor nominal, o mediante la inserción de un tornillo dieléctrico en el interior del soporte de la bobina, lo que varía la capacidad distribuida entre los conductores interior y exterior.

Los filtros con resonadores helicoidales se diseñan utilizando resonadores en cuarto de onda. Para el caso más frecuente de filtros de banda estrecha y alta selectividad se recurre a la caracterización del dispositivo mediante los coeficientes de acoplamiento entre resonadores y los decrementos de los resonadores extremos.

El acoplamiento entre resonadores es el problema más delicado del diseño. Se han estudiado tres tipos de acoplamiento: *1*) El acoplamiento realizado mediante una apertura en el blindaje, adyacente a los dos resonadores, enfrentando los extremos libres de ambos. Este acoplamiento se debe principalmente al campo eléctrico, puesto que se enfrentan los puntos de máxima

tensión. Por ello, se le denomina acoplamiento capacitivo. *2*) El acoplamiento realizado mediante una apertura en el blindaje, enfrentando los extremos cortocircuitados a masa. Este acoplamiento se debe principalmente al campo magnético, puesto que se enfrentan los puntos de máxima corriente. Por ello, se denomina acoplamiento inductivo. *3*) El acoplamiento se consigue disponiendo los resonadores sobre el mismo eje.

Se ha elaborado una teoría sobre el acoplamiento entre resonadores helicoidales con ejes paralelos y se han demostrado unas ecuaciones de diseño que permiten calcular las dimensiones mecánicas de la ventana de acoplamiento para obtener el coeficiente de acoplamiento deseado. En esa misma ilustración se presentan los datos experimentales y los calculados según la teoría elaborada para el acoplamiento entre dos resonadores cuyas características son: S = 3,175 centímetros; d = 2,1 centímetros; N = 4,3 vueltas; b = 3,75 centímetros; y Z_0 = 2185. Se estableció que, fijadas las características de los dos resonadores acoplados, el coeficiente de acoplamiento K_{12} prácticamente sólo depende de h/d, donde h es la altura de la ventana de acoplamiento y d su anchura, la cual se hace igual al diámetro de la hélice que forma el conductor interior del resonador.

Las medidas realizadas con diferentes resonadores muestran que la variación de log K_{12} en función de log (h/d) es prácticamente lineal. Con ello, desde un punto de vista práctico, se puede establecer que:

$$K_{12} = A\ (h/d)^{a},$$

donde A y a son dos coeficientes cuyos valores pueden calcularse mediante la teoría elaborada, en función de las características de los resonadores utilizados y de la frecuencia de trabajo.

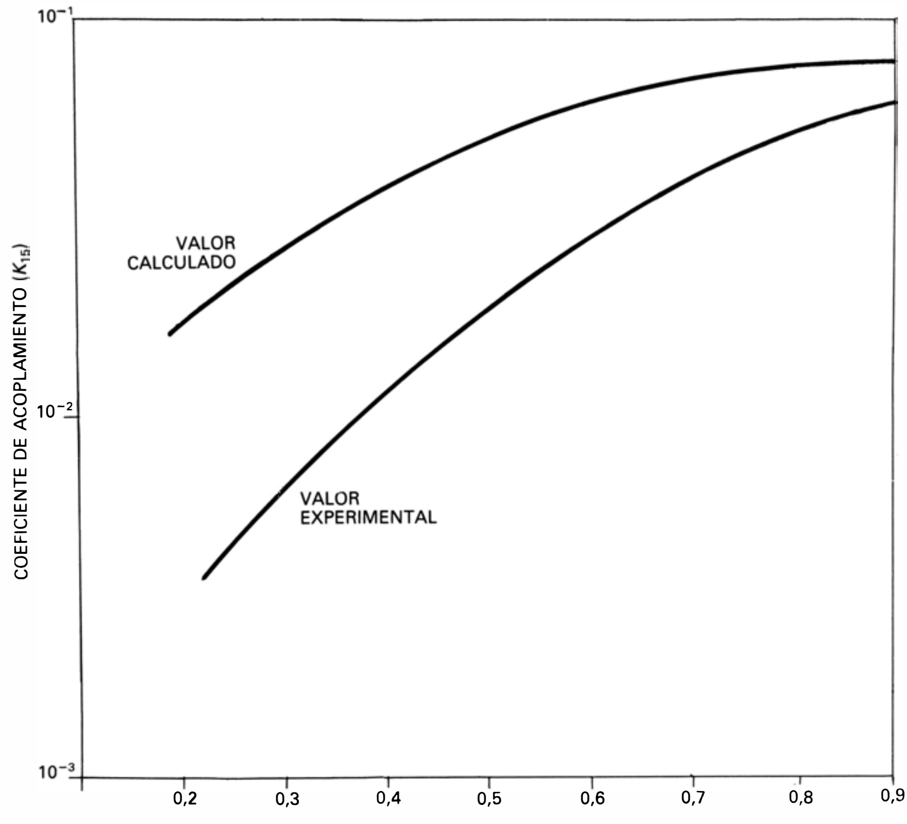
En vez de situar cada resonador en cavidades individuales, se pueden disponer a lo largo de un eje común como se muestra en la ilustración superior de la página siguiente. Este procedimiento presenta grandes ventajas en el proceso de fabricación y puede reducir el volumen y el peso del filtro. A pesar de su apariencia, hay que resaltar que los resonadores utilizados son equivalentes eléctricamente a los utilizados en la anterior estructura y sólo varía el procedimiento de acoplamiento. En efecto, el acoplamiento entre resonadores se debe, al mismo tiempo, a la inductancia mutua entre las hélices que forman los conductores exteriores y al acoplamiento capacitivo distribuido. El estudio teórico elaborado permitió establecer que, fijadas las dimensiones, el acoplamiento depende fundamentalmente de la separación p entre los extremos de las bobinas.

En la ilustración inferior de la página siguiente se muestra el coeficiente de acoplamiento teórico deducido, y los resultados prácticos obtenidos en el laboratorio para dos resonadores, en función de la separación p entre ellos.

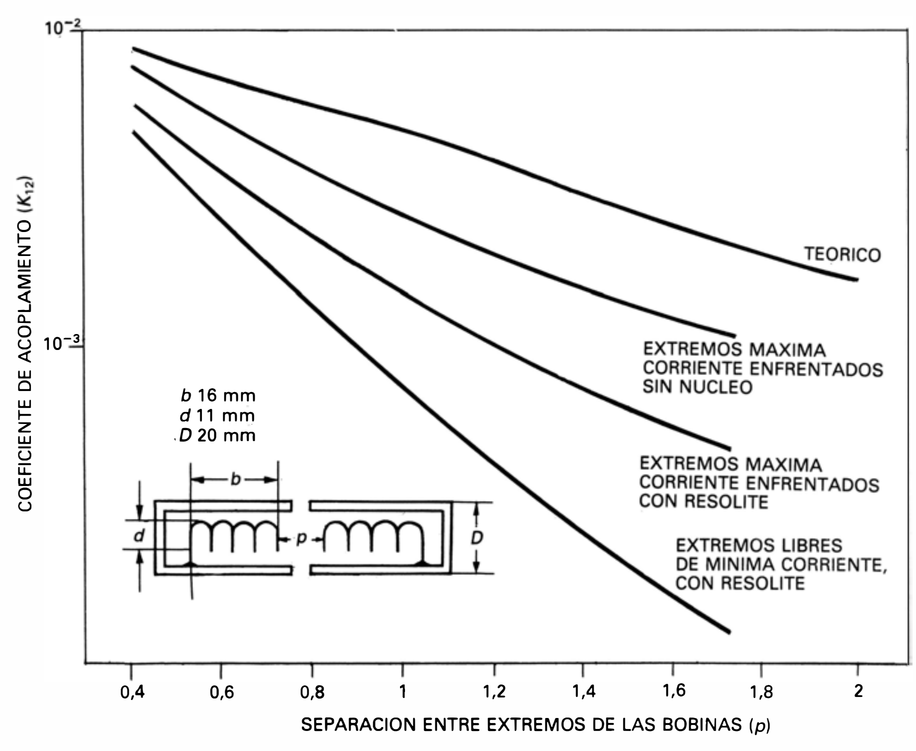
En ambos casos aparece la misma forma de variación con p, pero los valores prácticos están demasiado alejados de los teóricos. La curva teórica aparece como un límite superior, como era de esperar, ya que en el estudio teórico no se ha considerado el efecto capacitivo entre bobinas ni el producido por el soporte dieléctrico. La mayor aproximación se consigue para pequeñas se-

paraciones, donde la corriente es más uniforme y, en el caso de estar enfrentados los extremos de máxima densidad de corriente, sin soporte dieléctrico.

Los resultados de este trabajo fueron aplicados en la realización de los filtros de la unidad de radiofrecuencia del "Teléfono Supletorio Sin Hilos", diseñado en el Centro de Investigación y Estudios de la Compañía Telefónica Nacional de España. Este equipo es el único en su género a trabajar en la banda de UHF (450 megahertz) con canalización de 25 kilohertz y de acuerdo con las características preferidas por los organismos internacionales para equipos radioeléctricos móviles. (Antonio Castillo Holgado.)



Relación entre coeficiente de acoplamiento y separación entre resonadores



Valores calculados y experimentales para el acoplamiento entre dos resonadores

Archibacterias

Estas insólitas bacterias no son, en términos genealógicos, ni procariotas ni eucariotas. Tal descubrimiento significa que no hay dos líneas de descendencia evolutiva, sino tres: las archibacterias, las verdaderas bacterias y los eucariotas

Carl R. Woese

Los filósofos de la naturaleza mantenían, antiguamente, que la vida sobre la tierra era fundamentalmente dicotómica: todos los seres vivos eran animales o vegetales. Al descubrirse los microorganismos, se les dividió de la misma manera. Se consideraron animales los grandes y móviles, en tanto que los que no mostraban movimiento, bacterias incluidas, se les supuso vegetales. A medida que fueron avanzando los conocimientos sobre el mundo microscópico, se puso de manifiesto que no bastaba una clasificación de dos ramas; por ello se fueron introduciendo categorías adicionales: hongos, protozoos y bacterias. Andando el tiempo, se asistió a una nueva simplificación. Parecía que, a pesar de todo, la vida podía ser dicotómica, aunque a un nivel más profundo, es decir, a nivel de la estructura de la célula viva. Todas las células parecían pertenecer a uno de estos dos grupos: los eucariotas, células con un núcleo bien formado, y los procariotas, que no poseen tal núcleo. Las plantas y los animales pluricelulares son eucariotas, así como muchos organismos unicelulares. Los únicos procariotas son las bacterias (incluidas las cianobacterias, que se denominaron anteriormente algas verde-azuladas).

En los últimos años, mis colegas y yo nos hemos visto obligados a proponer una revisión fundamental de este cuadro. Entre las bacterias, hemos encontrado un grupo de organismos que no parecen pertenecer a ninguna de las dos categorías básicas. Los organismos que hemos estudiado son procariotas en el sentido de que no poseen un núcleo; su propio aspecto externo recuerda mucho el de las bacterias ordinarias. En su bioquímica, sin embargo, y en la estructura de ciertas macromoléculas, difieren tanto de los procariotas como de los eucariotas. Filogenéticamente no son ni lo uno ni lo otro. Constituyen un nuevo "reino primario", con una posición única en la historia de la vida.

A estos organismos los hemos denominado archibacterias. El término refleja una conjetura, sin comprobar, acerca de su situación en la escala evolutiva. Las pruebas filogenéticas sugieren que las archibacterias son tan antiguas, al menos, como los restantes grupos principales. Algunas archibacterias poseen, además, un metabolismo particularmente adecuado a las condiciones que se supone prevalecieron en los primeros tiempos de la vida sobre la tierra. De lo que se deduce que, quizás, este novísimo grupo de organismos sea, de hecho, el más antiguo.

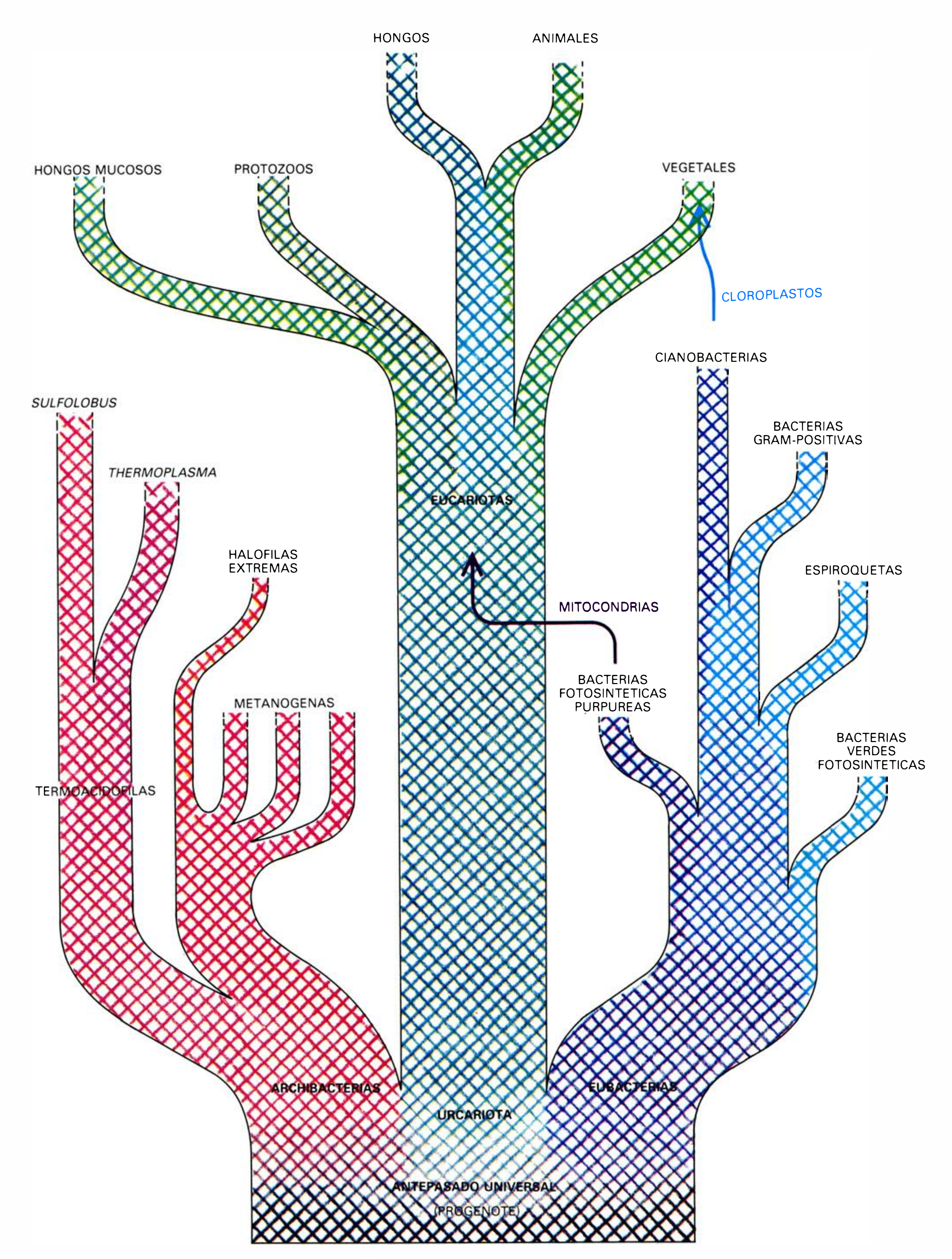
El registro evolutivo

La tierra tiene una edad de cuatro mil quinientos millones de años. Si atendemos a los restos microscópicos fósiles, se diría que ha estado habitada durante menos de una séptima parte de ese tiempo: la progresión evolutiva completa, desde las formas marinas más antiguas hasta el hombre, comprende solamente seiscientos millones de años. Las huellas fósiles de organismos celulares, demasiado pequeños para percibirlos a simple vista, nos cuentan una historia diferente. En particular, los microfósiles de bacterias abundan en los sedimentos de todas las edades; se han hallado en las rocas sedimentarias intactas más antiguas que se conocen, depósitos australianos que datan de hace tres mil quinientos millones de años. A lo largo de un gran lapso de tiempo, durante el cual no existieron formas superiores, surgieron las bacterias y se irradiaron hasta formar una amplia variedad de tipos, que habitaron en múltiples nichos ecológicos. Esta era de los microorganismos constituye el período más importante de la historia evolutiva, no sólo por su duración, sino también por la naturaleza de los acontecimientos evolutivos que ocurrieron a lo largo de estos miles de millones de años.

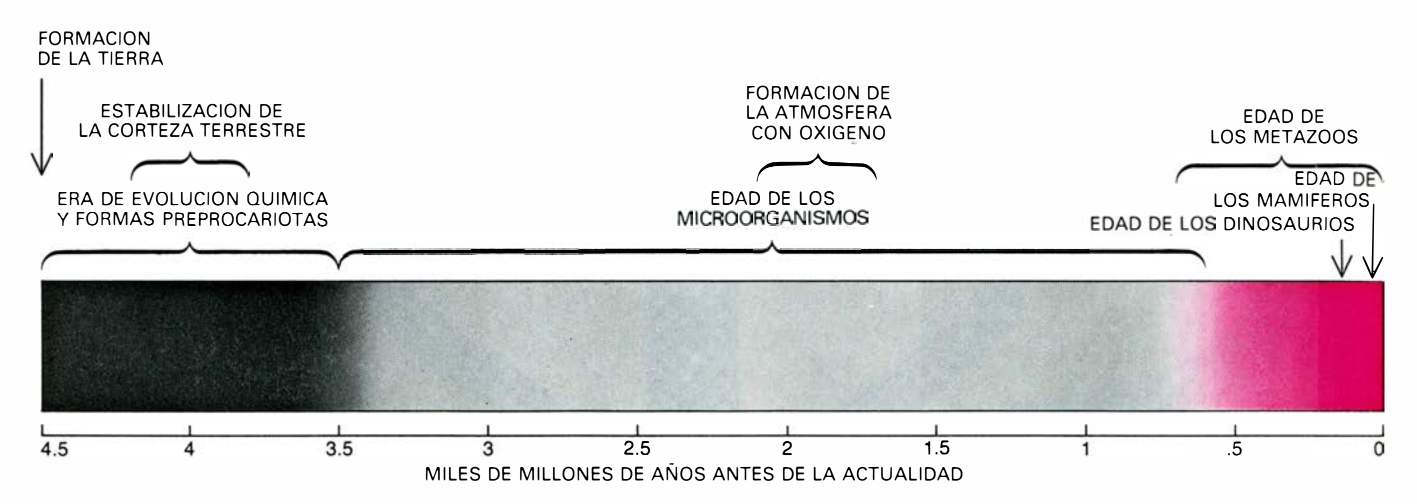
Hasta hace poco, sin embargo, no se sabía casi nada acerca de la edad de los microorganismos. Los microfósiles bacterianos no son estructuras muy informativas; es poco lo que puede inferirse de la huella de una pequeña esfera o de una varilla minúscula. Las principales indicaciones paleontológicas acerca de la naturaleza de las primitivas bacterias no han surgido de los microfósiles individuales, sino de estructuras macroscópicas denominadas estromatolitos, que se cree que son capas de bacterias fosilizadas: colonias de bacterias incrustadas en minerales. Hoy día, tales estructuras están formadas principalmente por varias clases de bacterias fotosintéticas, generalmente cianobacterias. Los estromatolitos de reciente fosilización se parecen tanto a los antiguos, que resulta razonable pensar que las antiguas estructuras fueron construidas también por bacterias fotosintéticas. Algunas, al menos, de las bacterias antiguas debieron haber sido, pues, fotosintetizadoras. Aparte de este único hecho, nada podía establecerse con certeza acerca de los primeros microorganismos. Todo el árbol evolutivo de las bacterias permanecía oscuro, así como la base del árbol de las formas superiores de vida.

Mas, para reconstruir los acontecimientos evolutivos primitivos, los biólogos no se hallan circunscritos a los datos geológicos fósiles. La propia célula encierra claves de su pasado en las secuencias de aminoácidos de sus proteínas y en las secuencias de nucleótidos de sus ácidos nucleicos, el ADN y el ARN. Este dato vivo es, en potencia, mucho más rico y amplio que el registro fósil, y se remonta en el tiempo mucho más atrás que los fósiles más viejos; alcanza hasta el período en el que existió el antepasado común de toda la vida.

A la lectura del registro bioquímico hubo de preceder una tecnología que permitiese determinar (al menos en parte) la secuencia de un gen, o del



TRES REINOS PRIMARIOS propone el autor para explicar el descubrimiento de que las archibacterias difieren fundamentalmente de las demás bacterias, denominadas eubacterias o bacterias verdaderas. Ambas coinciden en ser procariotas: células sencillas, carentes de núcleo, que difieren, en sus propiedades estructurales, de las células eucariotas, dotadas de núcleo y otros orgánulos subcelulares. Desde el punto de vista genealógico, sin embargo, las archibacterias y las eubacterias no guardan más relación entre sí que la que tienen respectivamente con los eucariotas. Para el autor, las archibacterias, las eubacterias y un urcariota –la célula eucariota original– derivaron de un antepasado común (el progenote), mucho más sencillo. Los eucariotas evolucionaron cuando el urcariota sirvió de "huésped" de endosimbiontes bacterianos, que, por evolución, darían lugar a las mitocondrias y los cloroplastos.

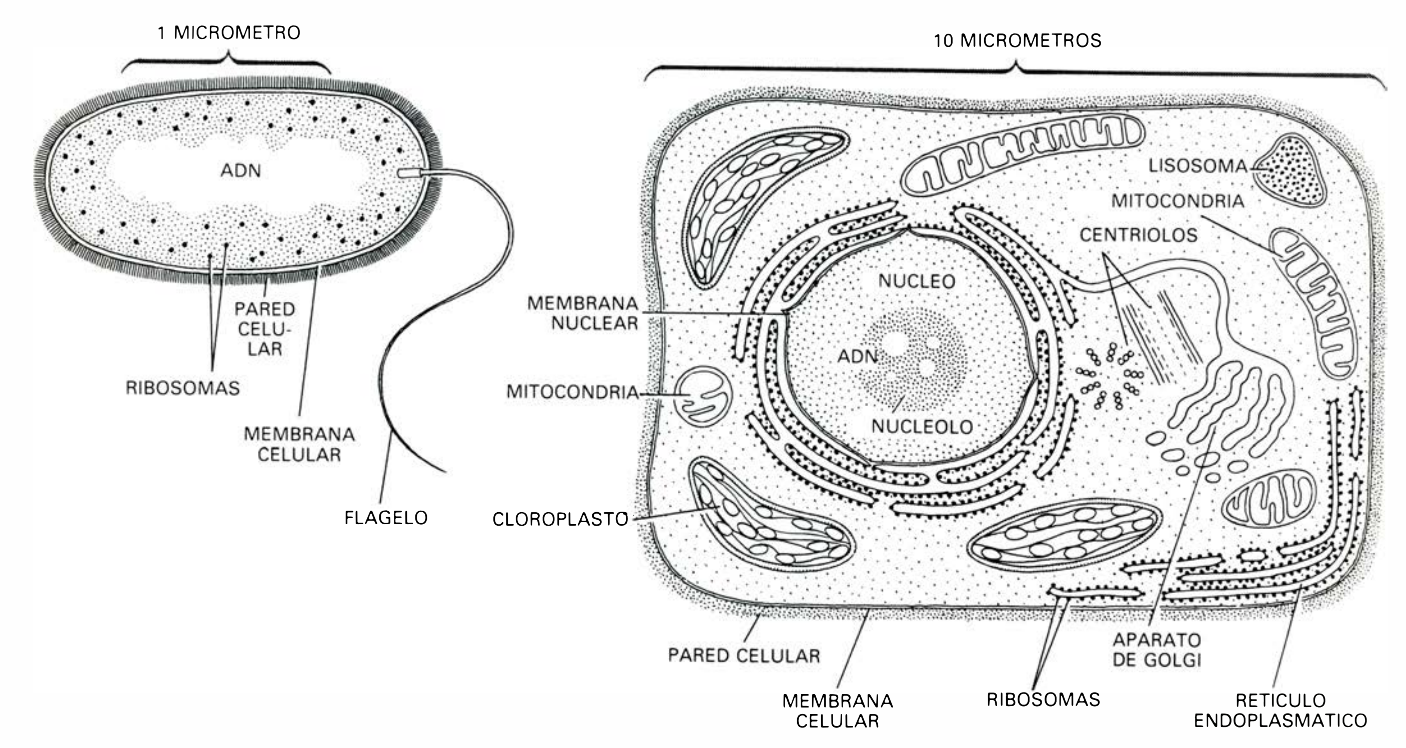


LA ERA DE LOS MICROORGANISMOS, que duró unos tres mil millones de años, abarca la mayor parte de la evolución biológica. Se han encontrado microfósiles de células procariotas en depósitos de unos tres mil quinientos millones de años de antigüedad. A estas células las debieron preceder otras más sencillas. Los primeros fósiles eucarióticos sólo tienen unos mil trescientos millones de años. Poco es lo que se sabe del proceso evolutivo durante la era de los microorganismos. Los restos fósiles microscópicos se remontan sólo hasta unos 600 millones de años, por la época de los primitivos metazoos.

ARN o la proteína codificados por el gen. Desde hace unos 25 años ello es factible en el caso de las proteínas; ahora bien, la secuenciación directa de ADN y de ARN se conoce sólo desde hace unos cinco años. La nueva tecnología para secuenciar ácidos nucleicos debería permitirle a los biólogos avanzar, en un tiempo relativamente corto, mucho más en la historia de la vida de lo que se había sospechado. Aplicando técnicas de secuenciación al secular problema de las relaciones naturales entre las bacterias, mis colegas y yo llegamos a reconocer que las archibacterias constituían una tercera forma de vida.

Eucariotas y procariotas

Para apreciar la posición especial de las archibacterias resulta útil considerar algunas de las características que definen a los eucariotas y los procariotas. La célula eucariótica es relativamente grande: mide unos 10 micrometros de lado. Rodeada por una doble membrana, distinguimos en su interior cierto número de estructuras, definidas a su vez por membranas. El núcleo contiene la mayor parte del material genético de la célula. En las mitocondrias, con forma de varilla, se produce la respiración, que genera la principal moneda energética de la célula, el trifosfato de adenosina (ATP). En las células vege-



LOS PROCARIOTAS Y LOS EUCARIOTAS presentan diferencias fundamentales a nivel estructural, como se aprecia en estos dibujos esquemáticos de una célula procariota (*a la izquierda*) y de una célula eucariota (*a la derecha*). El procariota es, con diferencia, la célula más pequeña. Se distinguen pocas estructuras subcelulares, incluso a la escala que alcanza el microscopio electrónico; suelto en el citoplasma hay un filamento único de material genético, ADN. Tanto las archibacterias como las eubacterias son procariotas y comparten propiedades estructurales propias de este nivel de organización. La célula eucariota es mucho mayor y posee numerosas estructuras subcelulares. Su ADN, que constituye un complejo con proteína, se organiza en forma de cromosomas, dentro de un núcleo rodeado por una membrana. Las mitocondrias llevan a cabo la respiración celular; en una célula vegetal hay cloroplastos, que dirigen la fotosíntesis. El aparato de Golgi es un orgánulo secretor; el retículo endoplasmático es un sistema de membranas, sobre el que, en algunas zonas, se disponen los ribosomas (en éstos se traduce a proteína la información genética). Eucariotas son las células más complejas que las bacterias.

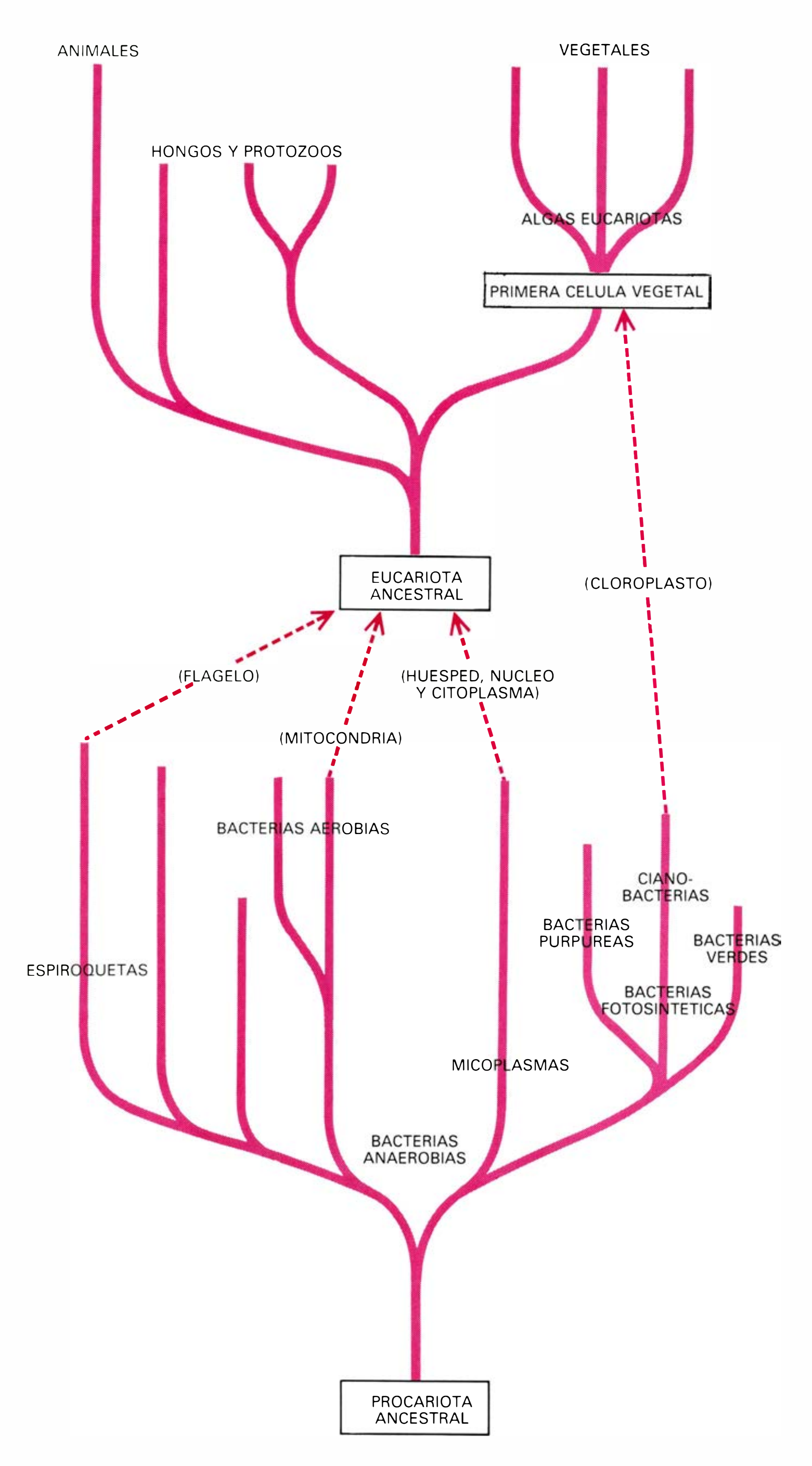
tales, el cloroplasto, otro cuerpo de forma alargada, convierte la energía de la luz en la energía química del ATP. Suele haber otras estructuras especializadas, tales como el aparato de Golgi (para la secreción) y los cilios (para la motilidad). Muchas células eucariotas están guarnecidas con un sistema de membranas, el retículo endoplasmático, que proporciona una superficie sobre la cual tienen lugar importantes reacciones, tales como la síntesis de proteínas.

Muy distinta es la célula procariota. Sus magnitudes, comparadas con las de la eucariota, son mucho más reducidas: unas diez veces menor en medidas lineales y, por tanto, unas 1000 veces menor en volumen. También la célula procariota está circunscrita por una doble membrana y suele poseer, además, una pared celular rígida. Carece, por contra, de las estructuras características de la célula eucariota; no hay mitocondrias, ni cloroplastos, ni, por supuesto, un núcleo rodeado por una membrana. El genoma, el complemento total de material genético, se limita a unos 2000 o 3000 genes en la célula procariota; el genoma eucariota típico supera esa cifra en varios órdenes de magnitud.

La distinción entre eucariotas y procariotas se definió inicialmente en términos de estructuras subcelulares apreciables al microscopio. A ese nivel, todas las células parecían ser, o grandes y complejas y, por tanto, eucariotas, o bien pequeñas y sencillas y, por ende, procariotas. La distinción entre los dos tipos de células se llevó hasta el nivel biológico básico: el molecular. Aquí, las células eucariotas y las procariotas comparten muchas características. Por ejemplo, traducen información genética a proteínas de acuerdo con el mismo código genético. Aun cuando los procesos moleculares sean los mismos, difieren los pormenores de cada forma: son característicamente eucarióticos o característicamente procarióticos. Por ejemplo, las secuencias de aminoácidos de varios enzimas tienden a ser típicamente procarióticas o eucarióticas. Todas estas diferencias entre grupos y las similitudes dentro de cada grupo hicieron creer a la mayoría de los biólogos que el árbol de la vida tenía dos troncos principales, un tronco procariota y otro eucariota.

La conclusión se sacó con apresurada ligereza, tal era el atractivo estético de la dicotomía. Que a nivel microscópico haya dos tipos de células no justifica que sólo existan dos tipos a nivel molecular.

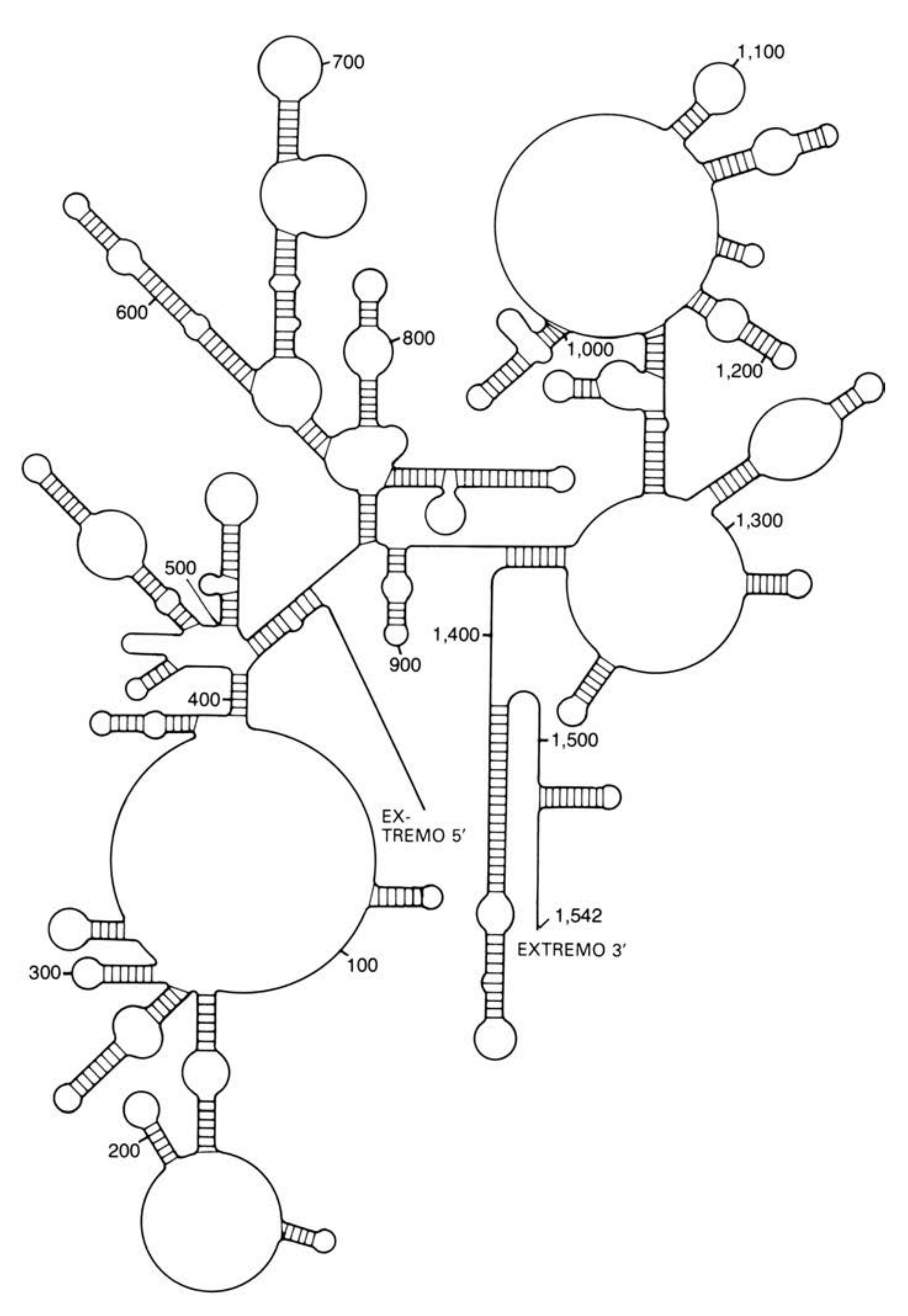
La verdad es que la relación evoluti-



EL ARBOL DE LA VIDA, según se dibujaba antes del descubrimiento de las archibacterias. Tenía dos líneas primarias de descendencia: los procariotas y los eucariotas, estos últimos derivados de los anteriores. Se daba por supuesto que las primeras células habían sido bacterias anaerobias (procariotas) que obtenían energía por fermentación. Originaron varias sublíneas. Después de que la atmósfera se enriqueciera en oxígeno, ciertas células anaerobias que habían perdido su pared celular (micoplasmas) establecieron una relación endosimbiótica con bacterias más pequeñas que habían ingerido. Una bacteria endosimbiótica aerobia (que respira oxígeno) dio, por evolución, la mitocondria; una cianobacteria fotosintética, el cloroplasto y una espiroqueta (tal vez) se convirtió en flagelo (un órgano de la motilidad). Así fue como una célula eucariota ancestral evolucionó y dio origen, a su vez, a los protozoos, hongos, animales y plantas. El dibujo se ha hecho a partir de un esquema de Lynn Margulis, de la Universidad de Boston.

va entre procariotas y eucariotas es más compleja que lo que podría indicar la evidencia antes citada. Dos orgánulos eucariotas, las mitocondrias y los cloroplastos, contienen cada uno su propio ADN. Además, los pigmentos que hay en el cloroplasto (las clorofilas y los carotenoides) son similares a los que se encuentran en las cianobacterias. Mitocondrias y cloroplastos tienen el tamaño de las bacterias; su maquinaria para traducir información genética a proteínas difiere de la maquinaria propia de la célula eucariota y posee cierto número de propiedades en común con la de los procariotas.

Estos y otros hechos apuntalaron la idea de que las mitocondrias y los cloroplastos descendían de procariotas que quedaron atrapados dentro de una célula mayor y establecieron una relación endosimbiótica con ella. Se supone que la mitocondria fue una bacteria respiratoria y el cloroplasto un pariente fotosintetizador de las cianobacterias. Esta conjetura, que en su forma más sencilla data de hace un siglo, quedó probada, para el cloroplasto, al demostrarse que la secuencia de nucleótidos de una de las clases de ARN del orgánulo, el ARN ribosómico, estaba específicamente relacionada con las secuencias del ARN ribosómico de las cianobacterias. Análogamente, el ARN ribosómico de la mitocondria, en los vegetales, parece ser del tipo bacteriano. Por tanto, al menos dos líneas de descendencia procariotas parecen estar representadas en la célula eucariota.



ARN RIBOSOMICO 16*S*. Las secuencias de nucleótidos de esta molécula se han comparado en distintos organismos al objeto de establecer sus relaciones filogenéticas. La molécula es un componente del ribosoma, la máquina molecular que sintetiza proteínas; la designación 16*S* se refiere a la velocidad con que sedimenta la molécula en una centrífuga, medida en unidades Svedberg. El ARN constituye una larga cadena de subunidades, denominadas nucleótidos, cada una de las cuales se caracteriza por una de las cuatro bases siguientes: adenina (*A*), uracilo (*U*), guanina (*G*) y citosina (*C*). Las dos primeras bases y las dos últimas son complementarias: pueden unirse mediante enlaces de hidrógeno formando pares, de manera que *A* se empareja con *U* y *G* con *C*. El apareamiento de bases determina la estructura secundaria de la molécula, esto es, el plegamiento inicial de la molécula, del que resultan unas 50 estructuras bicatenarias cortas, en las que las bases están apareadas (*regiones rayadas*). La ilustración, de Alan D. Iselin, muestra la estructura secundaria del ARN 16*S* de la eubacteria *Escherichia coli*, cuya secuencia completa fue determinada por Harry F. Noller, Jr., de la Universidad de California en Santa Cruz.

El urcariota

La pregunta lógica que surge es: ¿de dónde procede el resto de la célula eucariota? ¿qué fue de la célula huésped original: el urcariota? Se admite, en general, que la porción mayor de la célula eucariota (el núcleo y las estructuras citoplasmáticas) representan una línea independiente de descendencia. No está claro cuál sea la naturaleza exacta de la célula ancestral. La mayoría de los investigadores tienden a creer que la línea eucariota principal surgió también de entre las bacterias ordinarias. El razonamiento viene a ser el siguiente: alguna bacteria anaerobia, que obtenía su energía por fermentación de nutrientes (y no por oxidación), perdió en algún momento su pared celular rígida. Se conocen organismos de este tipo: los micoplasmas. Una cepa de micoplasmas evolucionó luego adquiriendo la capacidad de englobar otros organismos, capacidad retenida hoy por muchos eucariotas. De entre las muchas clases de organismos que el micoplasma citado pudo ingerir, parece que dos establecieron una relación endosimbiótica estable con su huésped y se convirtieron, así, en la mitocondria y el cloroplasto. De esta manera, nació la célula eucariota. (El origen de su característica definitoria, el núcleo rodeado por una membrana, dista mucho de hallarse establecido.)

Aunque satisfactoria en algunos aspectos, esa argumentación no logra explicar las muchas diferencias que median entre procariotas y eucariotas. En particular, no explica las distintas características de los procesos moleculares o las grandes diferencias en las secuencias de aminoácidos de proteínas

funcionalmente análogas de una y otra clase de células. Con frecuencia se da por supuesto que las diferencias son mera consecuencia de los múltiples cambios menores en las configuraciones celulares por los que necesariamente habría tenido que pasar la célula, desde su sencilla condición de procariota hasta la más compleja de eucariota. Cuesta creer que todos estos cambios (cambios en la composición de casi todos los enzimas, por ejemplo) puedan explicarse razonablemente así.

Por ese motivo, algunos biólogos piensan que la línea de descendencia que dio lugar a las hipotéticas especies urcarióticas pudo haber divergido de la línea procariota en puntos precedentes, antes de que surgiera el propio antepasado de las bacterias. El urcariota podría haber evolucionado, por tanto, de un modo independiente hasta adquirir una forma de complejidad equiparable a las bacterias. Esa suposición proporcionaría, cuando menos, un intervalo más dilatado para la aparición de las diferencias entre procariotas y eucariotas. El urcariota, entonces, representaría una línea de descendencia distinta de la de los procariotas, de acuerdo con la dicotomía filogenética básica.

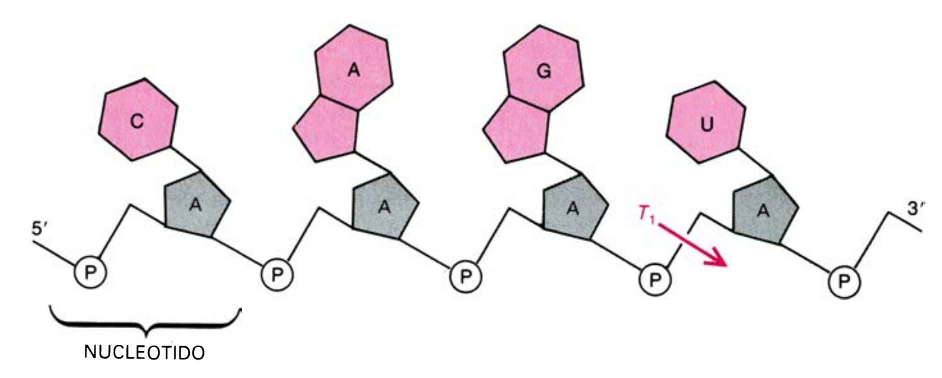
Así estaban las cosas a comienzos de la década de 1970. La filogenia de los eucariotas superiores, que cubría unos 500 millones de años, se comprendía razonablemente bien, salvo en lo referente al punto de unión de las principales ramas eucarióticas. Se disponía de una hipótesis definida, ampliamente aceptada, sobre la vía evolutiva de la célula eucariota. Las pruebas sobre la hipótesis y las respuestas a las restantes preguntas quedaban, sin embargo, en los puntos apartados y sin explorar de la filogenia bacteriana, en la era de los microorganismos.

Secuenciación genética

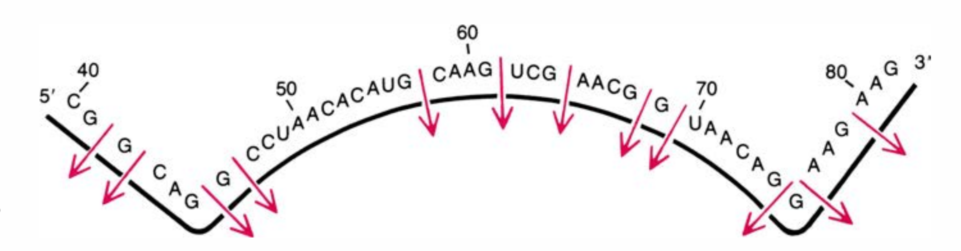
Las bacterias constituyen un mundo de extraordinaria variedad, mucho mayor de lo que revela el microscopio. Los nichos ecológicos en los que se encuentran exceden con mucho, en variedad, a los ocupados por las formas superiores de vida. Durante un siglo, los microbiólogos han tratado en vano de comprender las relaciones naturales entre las bacterias e imponer algún tipo de orden en la confusa distribución de formas, fisiologías y ecologías. La variedad entre bacterias es, principalmente, variedad dentro de la simplicidad; de ahí que no aporte mayor información acerca de las relaciones filogenéticas. En los organismos superiores, el ojo, por ejemplo, ha evolucionado cierto número de veces; se trata de un órgano suficientemente complicado como para que los distintos tipos que han evolucionado independientemente puedan distinguirse con facilidad. No suele ser éste el caso en lo que se refiere a la forma y la estructura de las bacterias; es fácil llegar, vía evolución, a las formas de varilla, esfera y espiral, que son las configuraciones bacterianas típicas. Lo mismo puede decirse de la bioquímica bacteriana. Aunque algunas características de estos microorganismos son indicadores filogenéticos válidos, resulta imposible decir por adelantado cuáles lo son y cuáles no.

Las secuencias genéticas de las células encierran su propio pasado. Los genes que se encuentran hoy en una célula son copia de los que existieron milenios atrás. No son una copia exacta, pues las mutaciones han alterado las secuencias genéticas originales, pero con frecuencia persisten vestigios de su primer estado. Lo que hace de un gen un registro soberbio de su pasado es su sencillez (es una disposición lineal) y el hecho de que el "espacio" de la secuencia genética sea enorme, de forma que, a lo largo de la extensión total de la evolución, sólo llega a materializarse una pequeña fracción de las posibles secuencias genéticas. Por ello, si dos genes son iguales a lo largo de un fragmento que comprende un número significativo de nucleótidos, puede aceptarse que poseen un antecesor común; se dice que estas moléculas, genéticamente relacionadas, son homólogas.

La secuencia genética suministra una triple información evolutiva: puede revelar relaciones genealógicas, ser una medida del tiempo de evolución y un registro de las características ancestrales. La magnitud de la relación entre dos genes para una misma función de diferentes organismos es igual a la magnitud de la relación entre dichos organismos. La magnitud de las diferencias entre las dos secuencias da una medida del tiempo transcurrido desde que los organismos divergieron a partir de un antepasado común. A partir de un amplio conjunto de secuencias relacionadas, puede construirse un árbol filogenético en el que los extremos de las ramas midan (aproximadamente) los tiempos relativos de las bifurcaciones. Finalmente, las comparaciones entre un amplio conjunto de secuencias homólogas permite reconstruir, con cierta precisión, diversas versiones ancestrales de un gen.



LOS NUCLEOTIDOS de una molécula de ARN constan de una base, un azúcar ribosa (*A*) y un grupo fosfato (*P*). Los enzimas ribonucleasas cortan la cadena por puntos específicos. La ribonucleasa T_1 la corta por hidrólisis (inserción de una molécula de agua) sobre el lado 3' del fosfato que sigue a un nucleótido con guanina. Por tanto, corta la molécula de ARN por varios puntos, dando lugar a cierto número de breves fragmentos, de uno o más nucleótidos, que terminan, todos ellos, por guanina (*G*).



EFECTO DE LA RIBONUCLEASA T_1 sobre un corto filamento de la molécula 16*S* de *E. coli*. Cuando un ARN 16*S* típico se corta de esta manera, su secuencia se divide en fragmentos ("palabras") cuya longitud oscila entre un nucleótido ("letra") y 20 nucleótidos. A continuación, se determina la secuencia de bases de cada una de las palabras. Las de seis letras o más se recopilan en un diccionario. Los diccionarios de dos organismos pueden compararse entonces en términos de un coeficiente de asociación, S_{AB}. El coeficiente es una fracción igual al doble del número de letras de las palabras (al menos con seis letras) comunes a los organismos *A* y *B*, dividido entre el número total de letras de tales palabras, en *A* y *B*.

Dado que la relación entre un gen y su producto (bien sea una proteína o uno o más tipos de moléculas de ARN) es generalmente colineal, la secuencia del producto es, de ordinario, tan útil para hacer estudios filogenéticos como la secuencia del propio gen. Hasta hace poco sólo podían secuenciarse las proteínas; de ahí que las primeras filogenias basadas en la evolución molecular se dedujeran a través de comparaciones entre proteínas. Los estudios comparativos de la proteína respiratoria citocromo *c* resultaron de un especial valor para confirmar y extender el árbol filogenético de los organismos superiores. Pero moléculas como esas del citocromo *c* no muestran la misma eficacia a la hora de establecer relaciones entre bacterias. Tales proteínas no gozan de una distribución universal; tampoco son estrictamente constantes en su función y, por ende, no sirven para un parangón absoluto. Por si fuera poco, a causa de la mayor antigüedad de los linajes bacterianos las diferencias entre secuencias pueden ser mucho mayores entre estos microorganismos que lo son entre los eucariotas. Estos factores hacen que las filogenias bacterianas, deducidas a partir de la evolución de las proteínas, sean incompletas e inciertas.

ARN ribosómico

Existen otros productos génicos que pueden servir de indicadores de las relaciones bacterianas. Todas las entidades autorreplicativas poseen necesariamente sistemas para mantener y propagar información genética, así como para traducirla a cadenas de aminoácidos, que constituyen las proteínas. La mayoría de las macromoléculas implicadas en estos procesos deben buscar su origen en los primeros estadios de la evolución de la célula; con certeza aparecieron antes de que las células se hiciesen suficientemente complejas como para denominarlas procariotas. Por tanto, podría esperarse que estas moléculas poseyeran propiedades de marcador filogenético.

La elección primera más razonable entre estas moléculas son las moléculas de ARN que forman complejos con proteínas y dan lugar a los ribosomas. En los ribosomas, la información genética se traduce a proteínas. El ARN ribosómico es fácil de aislar en cantidades manejables, pues una célula bacteriana típica posee entre 10.000 y 20.000 ribosomas. Además, las moléculas de ARN ribosómico parecen haber permanecido constantes en su función a lo largo de grandes distancias evolutivas. Merece destacarse ese dato, porque los cambios funcionales de una molécula acarrean cambios en la secuencia, que hacen difícil, sino imposible, comparar una secuencia molecular con otra y deducir relaciones filogenéticas. Otra ventaja más del ARN ribosómico estriba en que algunas porciones de sus secuencias, al menos, cambian con la suficiente lentitud como para que la secuencia ancestral común no quede totalmente borrada. En otras palabras, las secuencias permiten detectar las relaciones filogenéticas más profundas.

Existen tres clases de moléculas de ARN ribosómico. En las bacterias, el ARN ribosómico "grande" es el 23*S* (*S* significa unidades Svedberg, una medida de la velocidad de sedimentación en una ultracentrífuga, esto es, una medida indirecta del tamaño de la molécula); posee una longitud aproximada de 2900 nucleótidos. El "pequeño", designado por ARN ribosómico 16*S*, tiene una longitud de unos 1540 nucleótidos. Uno muy pequeño, el 5*S*, mide sólo 120 nucleótidos. Estos tamaños son similares en las células eucariotas: 18*S*, 25-28*S* y 5*S*. Podría pensarse que, por su facilidad de caracterización, el pequeño ARN 5*S* sería el más adecuado para los estudios filogenéticos. La realidad es que, como indicador de relaciones filogenéticas, no muestra la precisión de los ARN ribosómicos mayores, principalmente por razones estadísticas. (El ARN 5*S* presenta a veces grandes diferencias anómalas en su secuencia, según sea de una u otra especie.) Se opta por el ARN ribosómico 16*S* porque la molécula 23*S* es casi dos veces mayor y más de dos veces más difícil de caracterizar.

En la Universidad de Illinois, en 1969, decidí explorar las relaciones bac-

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	*Saccharomyces cerevisiae*	—	.29	.33	.05	.06	.08	.09	.11	.08	.11	.11	.08	.08	.10	.07	.08
2	*Lemna minor*	.29	—	.36	.10	.05	.06	.10	.09	.11	.10	.10	.13	.07	.09	.07	.09
3	Célula *L*	.33	.36	—	.06	.06	.07	.07	.09	.06	.10	.10	.09	.07	.11	.06	.07
4	*Escherichia coli*	.05	.10	.06	—	.24	.25	.28	.26	.21	.11	.12	.07	.12	.07	.07	.09
5	*Chlorobium vibrioforme*	.06	.05	.06	.24	—	.22	.22	.20	.19	.06	.07	.06	.09	.07	.05	.07
6	*Bacillus firmus*	.08	.06	.07	.25	.22	—	.34	.26	.20	.11	.13	.06	.12	.10	.07	.09
7	*Corynebacterium diphtheriae*	.09	.10	.07	.28	.22	.34	—	.23	.21	.12	.12	.09	.10	.10	.06	.09
8	*Aphanocapsa*	.11	.09	.09	.26	.20	.26	.23	—	.31	.11	.11	.10	.10	.13	.10	.10
9	Cloroplasto (*Lemna*)	.08	.11	.06	.21	.19	.20	.21	.31	—	.14	.12	.10	.12	.12	.06	.07
10	*Methanobacterium thermoautotrophicum*	.11	.10	.10	.11	.06	.11	.12	.11	.14	—	.51	.25	.30	.34	.17	.19
11	*Methanobrevibacter ruminantium*	.11	.10	.10	.12	.07	.13	.12	.11	.12	.51	—	.25	.24	.31	.15	.20
12	*Methanogenium cariaci*	.08	.13	.09	.07	.06	.06	.09	.10	.10	.25	.25	—	.32	.29	.13	.21
13	*Methanosarcina barkeri*	.08	.07	.07	.12	.09	.12	.10	.10	.12	.30	.24	.32	—	.28	.16	.23
14	*Halobacterium halobium*	.10	.09	.11	.07	.07	.10	.10	.13	.12	.34	.31	.29	.28	—	.19	.23
15	*Sulfolobus acidocaldarius*	.07	.07	.06	.07	.05	.07	.06	.10	.06	.17	.15	.13	.16	.19	—	.13
16	*Thermoplasma acidophilum*	.08	.09	.07	.09	.07	.09	.09	.10	.07	.19	.20	.21	.23	.23	.13	—

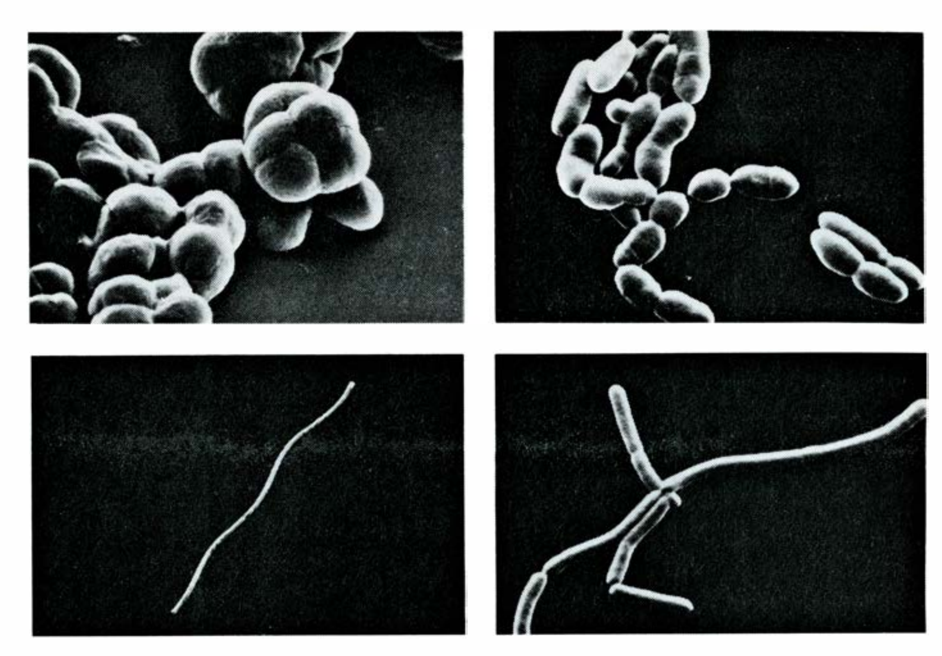
MATRIZ DE COEFICIENTES DE ASOCIACION para poner de manifiesto el grado de relación entre dos organismos cualesquiera. Cuanto más elevada sea la fracción S_{AB}, más estrecha será la relación. La distribución de las relaciones es significativa. Los eucariotas (*1-3*), las eubacterias (*4-9*) y las archibacterias (*10-16*) forman grupos distintos (*color*). Las archibacterias no están más relacionadas con las eubacterias que con los eucariotas. *Saccharomyces* es una levadura; *Lemna*, una planta acuática; las células *L* son una línea de células de ratón. El cloroplasto desciende de una cianobacteria.

terianas, comparando las secuencias del ARN ribosómico 16*S* de diferentes especies. Todavía no era posible entonces determinar la secuencia de nucleótidos de la molécula entera. Sí permitía la técnica secuenciar cortos segmentos de la molécula. Los enzimas ribonucleasas proporcionaban breves fragmentos de ARN, al cortar una cadena de ARN por sitios específicos. Los nucleótidos del ARN constan de un azúcar, la ribosa, un grupo fosfato y una de las siguientes cuatro bases nitrogenadas: adenina (*A*), uracilo (*U*), guanina (*G*) o citosina (*C*). El enzima ribonucleasa T_1 corta la cadena de ARN por un determinado enlace a un lado de cada nucleótido que lleve guanina como base. El enzima T_1 digiere, por tanto, un "texto" de ARN dando origen a cortas "palabras", llamadas oligonucleótidos. Los oligonucleótidos incluyen, y terminan por, una sola *G*, tal como *AACUCG* o *UCCUAUCG*.

Diccionarios de ARN

Los oligonucleótidos producidos así eran lo bastante cortos como para que pudieran secuenciarse utilizando las técnicas disponibles. Las palabras más pequeñas valen poco, pues se repiten muchas veces en cada molécula. Sin embargo, cuando la longitud de la palabra alcanza seis nucleótidos, no es probable que una determinada secuencia aparezca más de una vez en una molécula de ARN 16*S*. (Dada la constante *G* terminal, existen 3^5, esto es, 243 posibles secuencias de seis letras; una molécula típica de ARN 16*S* tiene aproximadamente unas 25 palabras de este tipo.) Cuando ARN 16*S* procedentes de diferentes organismos incluyen la misma secuencia de seis letras, puede considerarse, casi con seguridad, que se trata de una verdadera homología. Centrando la atención en palabras de seis o más letras, puede elaborarse un "diccionario" característico de un determinado organismo, que puede fácilmente compararse con otro de estos diccionarios y determinar relaciones genealógicas.

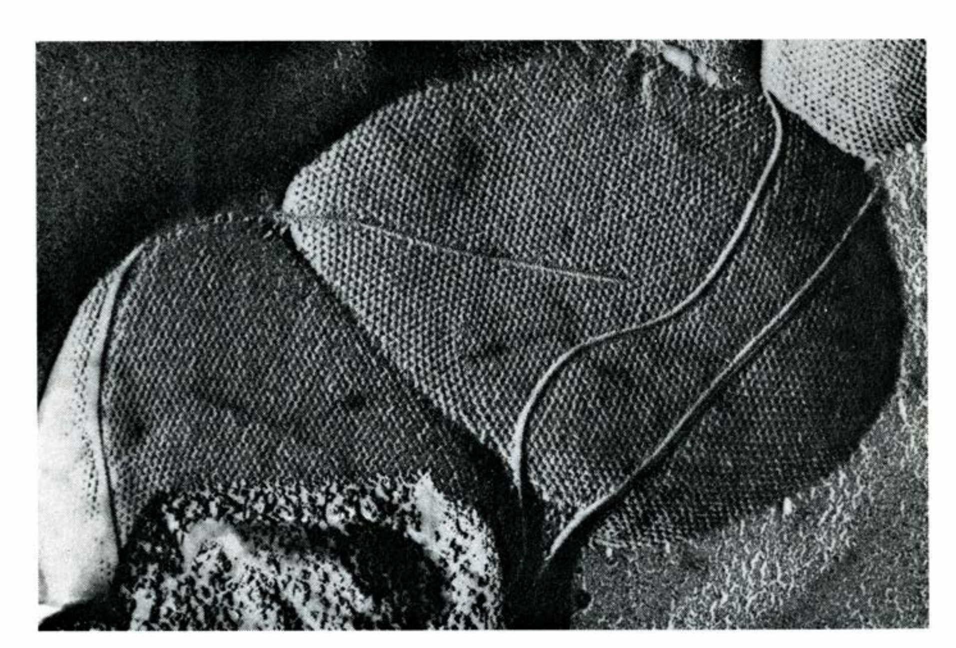
Una manera sencilla de analizar los datos es en términos de un coeficiente de asociación S_{AB}, que se define como el doble del número de nucleótidos que hay en las palabras comunes a ambos en los diccionarios *A* y *B*, dividido por el número de nucleótidos que hay en todas las palabras de los dos diccionarios. S_{AB} oscila entre 1, cuando los diccionarios *A* y *B* son idénticos, y menos de 0,1, cuando no están relacionados.



LAS METANOGENAS, bacterias anaerobias que generan metano (CH_4) a partir de hidrógeno y de dióxido de carbono, constituyen el grupo más extenso de archibacterias identificado hasta ahora. Cuatro géneros de metanógenas, que difieren ampliamente en tamaño y morfología, figuran en estas fotografías realizadas con un microscopio de barrido por Alexander J. B. Zehnder, del Instituto Federal Suizo de Tecnología. Se trata de *Methanosarcina* (*arriba, a la izquierda*), *Methanobrevibacter* (*arriba, a la derecha*), *Methanospirillum* (*abajo, a la izquierda*) y *Methanobacterium* (*abajo, a la derecha*), aumentadas, respectivamente, 2500, 5000, 1000 y 5000 veces. Las metanógenas sólo viven en ambientes libres de oxígeno.

(El coeficiente es generalmente mayor de cero, incluso en secuencias sin relación, debido a las correspondencias al azar.) Recopilando en una matriz los valores de S_{AB} para cierto número de organismos, puede distinguirse un patrón de relación o de falta de relación entre estos organismos. Además, mediante métodos estadísticos, es posible construir directamente, a partir de un conjunto de valores de S_{AB} para un grupo de organismos, un dendrograma, o árbol, que muestre las relaciones entre los miembros del grupo.

Hasta la fecha se ha caracterizado ya el ARN ribosómico de casi 200 especies de bacterias y eucariotas. La mayoría de las bacterias forman un grupo



INUSITADA PARED CELULAR DE *Methanogenium marisnigri*, una metanógena del fondo del Mar Negro, que aparece ampliada 70.000 veces en la micrografía electrónica de barrido obtenida por Frank Mayer, de la Universidad de Göttingen. El patrón en mosaico de las subunidades proteináceas es característico de varias archibacterias. Difiere de la típica pared celular eubacteriana, formada por subunidades de peptidoglicano, un polimero de glúcidos, del que carecen las paredes celulares de las archibacterias.

coherente, pero muy amplio (lo que equivale a decir antiguo). Se trata de las eubacterias, o bacterias verdaderas, que, como era de esperar, difieren bastante de los eucariotas. Las relaciones entre los diversos géneros (representados por las ramificaciones del árbol, determinadas a través del análisis del ARN ribosómico) contradicen muchas hipótesis sobre las relaciones bacterianas. Importa poner de relieve que las eubacterias están divididas en cierto número de ramas principales y que varias de las ramas incluyen a las bacterias fotosintéticas. Este hallazgo sugiere que todas las eubacterias derivan de un antepasado fotosintético común.

El descubrimiento de las archibacterias

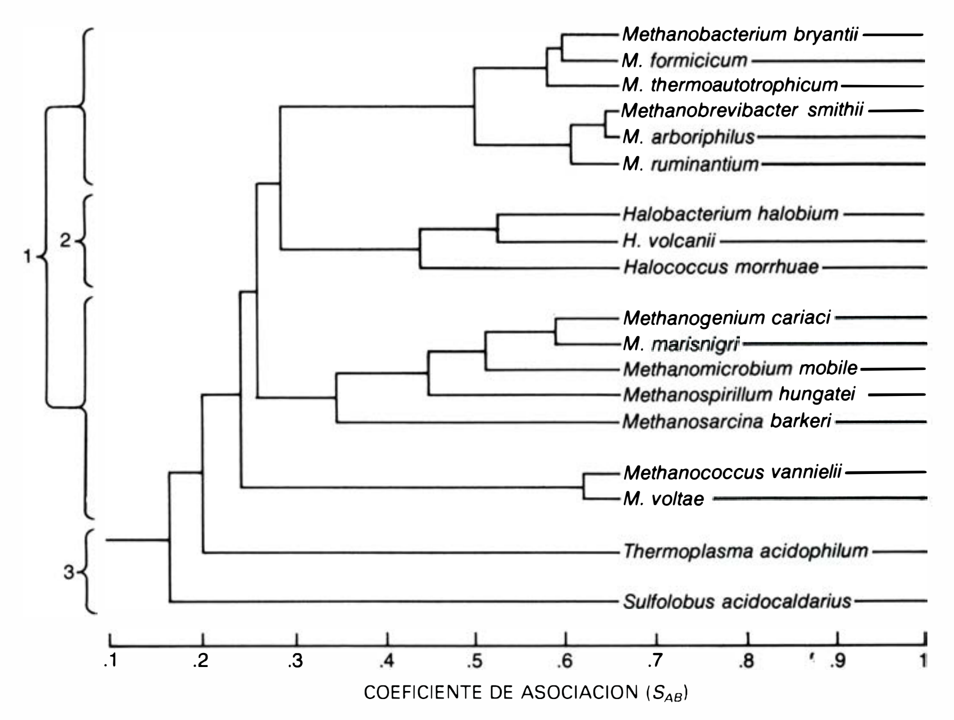
Con el progresivo estudio de las distintas bacterias saltó la sorpresa. En colaboración con Ralph S. Wolfe, observé el ARN ribosómico de las bacterias metanógenas. Estos inusitados organismos viven solamente en ambientes libres de oxígeno y generan metano (CH_4) mediante la reducción del dióxido de carbono (CO_2). Descubrimos que las metanógenas no encajaban en el grupo filogenético definido por las otras bacterias. Ciertamente, parecían representar una ramificación evolutiva que precedió en mucho al antepasado común de las bacterias verdaderas. El grupo de las metanógenas parecía ser, filogenéticamente, tan profundo –esto es, tan antiguo– como el grupo definido por las eubacterias.

No cabe duda de que las metanógenas y sus parientes son bacterias. Poseen el tamaño de éstas, carecen de membrana nuclear, tienen un bajo contenido de ADN, etcétera. Era de esperar, pues, que guardaran una relación más estrecha con otras bacterias que con los eucariotas. Nuestros análisis mostraron que no era así. Las metanógenas estaban relacionadas, por igual, con los eucariotas y con las eubacterias.

¿Cómo explicarlo? Se había supuesto la existencia de dos únicas líneas primarias de descendencia: los eucariotas y los procariotas. Teníamos ante nosotros un nuevo grupo de organismos: los metanógenos y sus parientes, que, en conjunto, se han denominado archibacterias. Su apariencia externa era, obviamente, la de las demás bacterias, por lo que se había supuesto que estaban en la línea de descendencia procariótica. No existen, sin embargo, notables diferencias en las características morfológicas que distingan filogenéticamente los procariotas de los eucariotas. Se trata de diferencias más sutiles y más antiguas: en las secuencias moleculares y en detalles funcionales a nivel molecular. Por ello, no hay razón para que dos líneas procarióticas no puedan ser tan diferentes entre sí como lo son de la línea eucariótica.

La idea resultaba demasiado novedosa para que se la aceptara sin oposición. Algunos biólogos rechazaron tajantemente la noción de una "tercera forma de vida". ¿Cómo podía algo que parecía una bacteria no ser una bacteria y no estar siquiera relacionado con las bacterias? Con el tiempo, la sencillez de nuestro argumento y la acumulación de pruebas se impusieron. Aunque hay biólogos que cuestionan todavía nuestra interpretación, la idea de que las archibacterias representan un grupo independiente al más alto nivel recibe un apoyo cada vez más generalizado.

La supuesta antigüedad de las archibacterias continúa siendo una hipótesis sin probar, aunque plausible. El fenotipo metanogénico parece cubrir un espacio filogenético tan grande, o mayor, que el abarcado por cualquier otro fenotipo bacteriano equiparable. Ello implica que las metanógenas son tanto o más viejas que cualquier otro grupo bacteriano. Es más, el metabolismo metanógeno (la reducción del dióxido de carbono a metano) encaja perfectamente con la clase de atmósfera que se supone que tenía la tierra primitiva: una atmósfera rica en dióxido de carbono, con algo de hidrógeno y virtualmente sin oxígeno. El nombre de archibacterias alude al hecho de que estos organismos debieron dominar en la biosfera primitiva. Al cambiar las condiciones, la necesidad de un ambiente anaerobio confinó a las metanógenas a una gama limitada de nichos relativamente inaccesibles.



DENDROGRAMA ARCHIBACTERIANO; se obtiene esa configuración arborescente a partir de los valores de S_{AB} y muestra las relaciones filogenéticas entre miembros de este reino primario. La mayoría de ellos son metanógenas (*1*), que son anaerobias (organismos que sólo sobreviven en ausencia de oxígeno) y generan metano por reducción del dióxido de carbono o de algunas otras fuentes de carbono muy sencillas. Las halófilas extremas (*2*) son bacterias aerobias (que respiran oxígeno), que habitan sólo en ambientes con una elevada concentración salina. Las termoacidófilas (*3*) son bacterias aerobias que viven en ambientes muy calientes, altamente ácidos. Las metanógenas parecen ser un grupo muy antiguo, del que debieron surgir las halófilas; las archibacterias termoacidófilas pudieron originarse por otra vía.

Formas archibacterianas

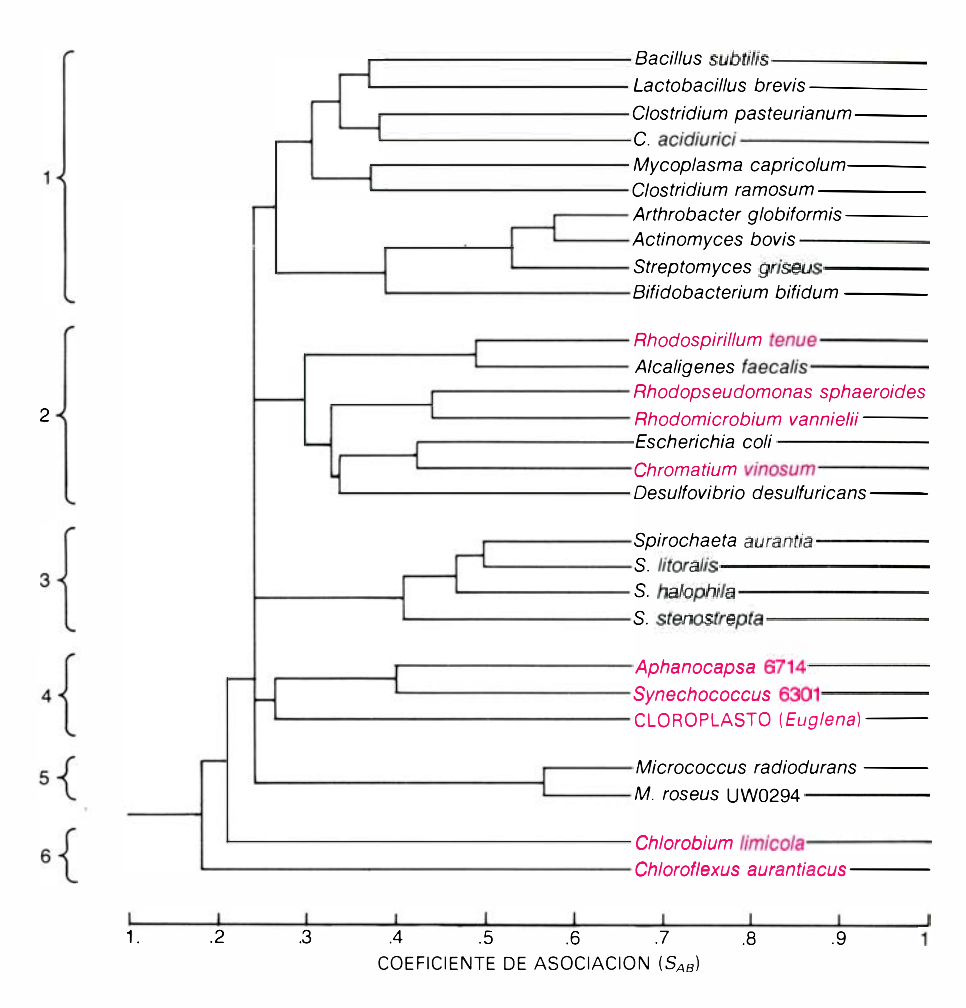
Las mediciones que revelaron la existencia de las archibacterias (diferencias en las secuencias del ARN) fueron de carácter genético y puramente cuantitativas. Nada descubrieron de la calidad de las diferencias –las diferencias fenotípicas– entre las archibacterias y las bacterias verdaderas. Si nuestra hipótesis de que las archibacterias constituyen un reino primario separado del de las bacterias verdaderas es correcta, una detallada inspección de las archibacterias debería mostrar que divergen tanto de las verdaderas bacterias, en su fenotipo molecular, como difiere cualquier grupo de las células eucariotas.

Las archibacterias son, ciertamente, organismos inusitados. Se sabe que el grupo comprende tres clases muy diferentes de bacterias: las metanógenas, las halófilas extremas y las termoacidófilas.

La forma dominante (en el sentido de que constituye una profunda agrupación filogenética) es la metanógena. Las bacterias que desprenden metano se conocen desde hace algún tiempo. Alessandro Volta descubrió, en 1776, que se generaba "aire combustible" en pantanos, corrientes y lagos cuyos sedimentos eran ricos en vegetación en descomposición, pero transcurrió largo tiempo antes de descubrirse que un microorganismo fuese el responsable de la generación del "gas de los pantanos". Las metanógenas, ampliamente distribuidas por la naturaleza, no existen al aire libre, porque el oxígeno las mata.

En tiempos remotos, las metanógenas debieron ser ubicuas. Hoy sólo viven donde se ha eliminado el oxígeno y disponen de dióxido de carbono e hidrógeno. Esto, generalmente, significa vivir en estrecha asociación con otras bacterias, tales como los clostridios, que metabolizan materia orgánica en descomposición y desprenden hidrógeno como producto de desecho. Las metanógenas se encuentran en aguas estancadas y en las plantas de tratamiento de aguas residuales de cloacas (en cantidades que han hecho posible su utilización comercial para obtener metano). Se encuentran también en la panza del ganado y de otros rumiantes y, en general, en el tracto intestinal de los animales. Las metanógenas pueden aislarse del fondo del océano y de los manantiales de aguas termales. A pesar de su intolerancia al oxígeno, gozan de una distribución universal.

Las halófilas extremas son bacterias que, para sobrevivir, requieren elevadas concentraciones de sal; algunas de ellas crecen fácilmente en salmuera saturada. Pueden conferir un color rojo a los estanques de evaporación donde se obtiene sal y decoloran y alteran el pescado salado. Las halófilas extremas crecen en hábitats salinos a lo largo del borde de los océanos y en ciertas aguas continentales, como las del Gran Lago Salado y el Mar Muerto. Aunque las halófilas extremas han sido ya objeto de estudio por parte de los microbiólogos desde hace tiempo, recientemente han cobrado nuevo interés. Por dos razones: mantienen fuertes gradientes en la concentración de ciertos iones a través de la membrana celular y utilizan esos gradientes para transportar diversas sustancias hacia dentro y hacia fuera de la célula. Además, las halófilas extremas poseen un mecanismo fotosintético relativamente sencillo, que no se basa en la clorofila, sino en un pigmento ligado a la membrana, la rodopsina bacteriana, muy parecido a uno de los pigmentos visuales.



EN ESTE DENDROGRAMA EUBACTERIANO se recogen seis grupos principales, tres de los cuales incluyen bacterias fotosintéticas (*color*); otros grupos quedan aún por definir. Las bacterias gram-positivas (*1*) poseen una gruesa pared celular con una composición única que absorbe y retiene el colorante de Gram. Las bacterias purpúreas fotosintéticas están agrupadas (*2*) con varios parientes, no fotosintetizadores, que presumiblemente han perdido una capacidad ancestral. Las espiroquetas (*3*) son largas bacterias de forma espiral. Las cianobacterias (*4*) son organismos fotosintéticos que producen oxígeno; el cloroplasto desciende de ellos. Algunas bacterias esféricas, con una pared celular atípica (*5*), destacan por su resistencia a la radiación. Finalmente, las bacterias fotosintéticas verdes (*6*) son también anaerobias.

Las termoacidófilas

El tercer tipo conocido de archibacterias es el termoacidófilo, cuyos miembros destacan también por su hábitat. *Sulfolobus*, uno de los dos géneros de termoacidófilas, se encuentra en los manantiales de aguas termales sulfurosas. Sus diversas especies crecen generalmente a temperaturas próximas a los 80 grados Celsius; hay variedades que se desarrollan a temperaturas superiores a los 90 grados C. Además, los manantiales donde medra *Sulfolobus* son extremadamente ácidos. Son comunes valores de *p*H inferiores a 2 (un *p*H de 7 es neutro). *Termoplasma*, el otro género de termoacidófilas, no se ha hallado hasta ahora más que en rescoldos de las pilas de carbón. Se trata de un micoplasma: sólo posee membrana celular, no pared celular.

Aunque los termoacidófilos archibacterianos se desarrollan, de modo exclusivo, en ambientes ácidos, el medio interior de la célula tiene un *p*H bastante moderado, próximo a la neutralidad; ello exige mantener un considerable gradiente de *p*H a través de la membrana de la célula. Al igual que en los halófilos extremos, el gradiente puede emplearse para bombear moléculas hacia dentro o fuera de la célula. Resulta interesante que, si se reduce la temperatura y, en consecuencia, *Sulfolobus* cesa de metabolizar, el *p*H interno de la célula no puede mantenerse próximo a la neutralidad, muriendo la célula.

Durante algún tiempo se admitió que

varios organismos, clasificados ahora como archibacterias, eran individuos algo peculiares. Se suponía siempre que semejante idiosincrasia debíase a una adaptación a un nicho peculiar o a un capricho bioquímico. Sin embargo, la medición filogenética del ARN ribosómico mostró que algunas de estas idiosincrasias, por lo menos, constituían características generales de un nuevo grupo de organismos. Ante la nueva situación, investigadores de muchos países se han aprestado a descubrir las propiedades generales que relacionan entre sí a las archibacterias y, una vez identificadas, averiguar si estas propiedades distinguen a las archibacterias de las otras dos formas principales, o las relacionan específicamente con una u otra de aquellas formas.

Característica general de las bacterias es que poseen una pared celular que contiene ácido murámico, un derivado azucarado, que es la base de un complejo polímero denominado peptidoglicano. Se sabía que un halófilo extremo y el termoacidófilo *Sulfolobus* se apartaban de esta regla general; se supuso, entonces, que poseían una estructura peculiar en su pared celular. Otto Kandler, de la Universidad de Munich, en colaboración con Wolfe, realizó un estudio sistemático de la estructura de la pared celular en otras archibacterias conocidas. Todas ellas resultaron ser atípicas. Las archibacterias tienen múltiples tipos de pared, pero ninguno de ellos es del tipo del peptidoglicano basado en el ácido murámico.

Lípidos y ARN

Se sabía también que la membrana celular de las halófilas extremas y de las termoacidófilas estaba formada por lípidos no habituales. Los lípidos, tanto de los eucariotas como de las eubacterias, constan principalmente de dos ácidos grasos de cadena recta, ligados por un extremo a una molécula de glicerol, vía enlaces de tipo éster (-CO-O-). Los lípidos de las halófilas extremas y de las termoacidófilas están también compuestos por un grupo glicerol ligado a dos largas cadenas hidrocarburadas, pero la conexión entre el glicerol y las cadenas es un enlace de tipo éter (-O-), no éster. Más aún, las cadenas hidrocarburadas no son rectas, sino ramificadas: cada cuatro átomos de carbono de la cadena hay un grupo metilo (CH_3). En otras palabras, el lípido archibacteriano básico es un diéter compuesto de glicerol y de dos moléculas de un alcohol, el fitanol. Después de examinar un cierto número de metanógenas en relación con su composición de lípidos, se confirmó nuestra suposición: sus lípidos eran éteres de glicerol ramificados, típicamente archibacterianos.

En el transcurso de los estudios sobre el ARN ribosómico surgió otra inesperada propiedad archibacteriana, que vino a ofrecer la primera clave del significado de las diferencias entre las archibacterias y las bacterias verdaderas. En el proceso de la traducción, desempeña un papel capital la molécula de ARN de transferencia. Reconoce secuencias de tres nucleótidos en el ARN mensajero que especifican determinados aminoácidos y aporta estos aminoácidos para su incorporación a la cadena de proteína. Un cierto número de nucleótidos de la molécula de ARN de transferencia están modificados, es decir, su estructura se ha alterado químicamente después de haberse incorporado a la molécula; normalmente, la modificación consiste en añadir un grupo metilo a la base o al azúcar del nucleótido. Los biólogos han deducido que existe una modificación característica en una cierta posición de casi todas las moléculas de ARN de transferencia: en esta posición, la base uracilo se ha metilado y forma timina (que normalmente se presenta sólo en el ADN, no en el ARN). Todos los ARN de transferencia de las archibacterias carecen de esta unidad de timina y, en su lugar, el uracilo se ha transformado, bien en una pseudopurina, bien en un nucleótido todavía por identificar.

Si comparamos el ARN ribosómico y el ARN de transferencia en los eucariotas, en las eubacterias y en las archibacterias, observamos un patrón general, del que el remplazamiento de la timina en los ARN de transferencia archibacterianos es sólo un ejemplo. En las tres líneas primarias de descendencia tienden a ser modificadas las mismas regiones del ARN, pero la naturaleza de la modificación varía de un reino a otro.

ENLACE ESTER
LIPIDO NORMAL
PALMITATO
GLICEROL
LIPIDO ARCHIBACTERIANO
ENLACE ETER
FITANOL
GLICEROL
CARBONO OXIGENO HIDROGENO

LOS LIPIDOS DE LA MEMBRANA de las archibacterias difieren de los que se encuentran en otros organismos. Los lípidos, tanto de las eubacterias como de los eucariotas, son ésteres de glicerol, de ácidos grasos de cadena recta, es decir, están compuestos de un alcohol de tres átomos de carbono, el glicerol, unido a cadenas de ácidos grasos, como el palmitato, mediante un enlace de tipo éster. Los lípidos archibacterianos, por el contrario, son diéteres en los que la unidad de glicerol se engarza mediante un enlace tipo éter a los fitanoles: cadenas ramificadas donde los átomos de carbono, a intervalos regulares, portan un grupo metilo (CH_3). Además, el glicerol tiene dos isómeros ópticos que se distinguen por la configuración de la molécula respecto al átomo central de carbono (*color*); los isómeros ópticos hacen girar la luz polarizada en sentidos opuestos. La configuración respecto al carbono central del glicerol que se encuentra en los lípidos archibacterianos es imagen especular de la configuración eubacteriana y eucariota.

Las diferencias son de dos clases. O la modificación de una determinada base es diferente en cada uno de los reinos, o bien se modifica en un reino una base determinada y en otro reino la modificación se realiza en una base adyacente. Tales formas de variación sugieren que las modificaciones han evolucionado por separado en cada una de las principales líneas de descendencia.

Se conocen otras diferencias moleculares entre las archibacterias y los demás grupos (por ejemplo, las que existen en la estructura de la subunidad del enzima ARN polimerasa), aunque la cifra no es muy alta. Pero no se trata tanto de que no existan más diferencias cuanto de que el mundo archibacteriano está todavía virgen. El estudio de la genética archibacteriana se halla en pañales; se han aislado muy pocos mutantes para su investigación. No se sabe nada todavía acerca de la expresión de los genes en las archibacterias, ni cuál sea la biología molecular fundamental de éstas. Ello no obstante, hasta donde ha podido avanzarse en su caracterización, se ha admitido que las archibacterias difieren significativamente de los otros dos grupos principales.

Una nueva perspectiva

El descubrimiento de un nuevo reino primario de organismos constituye un hallazgo de primera línea por derecho propio (comparable a salir al jardín y encontrarse con un organismo que no sea ni vegetal ni animal). Mas la importancia real del descubrimiento reside en lo que puede revelar acerca de la primitiva historia de la vida. Cuando se consideraban solamente dos líneas primarias de descendencia, no era fácil interpretar las diferencias existentes entre ambas. El reconocimiento de tres líneas de descendencia, equidistantes entre sí, proporciona una perspectiva mucho mejor para juzgar qué propiedades son ancestrales y cuáles han evolucionado recientemente. Con el descubrimiento de las archibacterias se hacen abordables, por tanto, dos problemas evolutivos centrales: la naturaleza del antepasado común de toda la vida y la evolución de la célula eucariota.

¿En qué estadio de la evolución de la célula tuvo lugar la división fundamental en reinos primarios? ¿Cuál era la naturaleza del antepasado universal? Se supone que el antepasado universal fue un procariota, la forma de vida más simple que existe hoy. Sin embargo, hubieron de precederle formas celulares más sencillas todavía. Aunque nada sabemos de ellas, cabe aventurar algunas de sus propiedades generales.

Analicemos el siguiente razonamiento. El mecanismo de traducción es complejo; comprende del orden de 100 grandes componentes moleculares. Es, además, sumamente preciso en su funcionamiento. Fabricar proteínas del tamaño común hoy día (cadenas de 100 a 500 aminoácidos), obteniendo un producto sin defectos el 90 por ciento de las veces o más, exige una tasa de error inferior a pocas partes por 10.000. Además, esta precisión debe alcanzarse mediante un mecanismo de dimensiones moleculares. Resulta imposible que este mecanismo haya podido evolucionar en un solo paso. La primitiva versión del mecanismo hubo de ser mucho más sencilla, menor y menos precisa.

La imprecisión en la traducción hubiera requerido la síntesis de proteínas que fuesen más cortas, y por ello menos específicas en su acción, que las proteínas actuales (de otra forma, la probabilidad de error al fabricar las proteínas hubiera sido demasiado grande). Entre las proteínas más pequeñas y menos específicas, se encontrarían los enzimas requeridos para elaborar la información genética. Si estos enzimas fueron menos precisos que las versiones actuales, la tasa de mutación de las células debió ser, necesariamente, más elevada y, el tamaño de su genoma, proporcionalmente menor. El proceso de traducción es la unión entre el genotipo y el fenotipo, entre la información y su expresión; a medida que el proceso evolucionaba para hacerse más preciso, la propia célula atravesó la correspondiente serie de refinamientos evolutivos. Evolucionó desde una entidad que poseía propiedades sencillas, funciones imprecisas y generales y una dotación de genes más bien pequeña, hasta una entidad que funcionaba con muchos enzimas altamente específicos y un complejo y exacto aparato genético. Con el fin de destacar los mecanismos genéticos y de traducción de las células primitivas y más sencillas, las denomino progenotes.

El progenote como antepasado

El descubrimiento de las archibacterias proporciona la perspectiva necesaria para abordar la cuestión de si el antepasado universal fue un procariota o un progenote. Aunque la cuestión dista mucho de estar zanjada, los primeros indicios permiten suponer que el antecesor universal fue, ciertamente, un progenote. En primer lugar, consideremos que las bacterias verdaderas y las archibacterias han existido, probablemente, al menos durante tres mil quinientos millones de años. El tiempo necesario para la evolución de las primeras bacterias verdaderas o archibacterias tiene que haber sido, por tanto, inferior a mil millones de años, y tal vez mucho menos. Sin embargo, los cambios evolutivos que han surgido dentro de cada uno de los reinos bacterianos, a lo largo del intervalo de tres mil millones de años o más, son mínimos si los comparamos con las diferencias que separan las archibacterias de las bacterias verdaderas, diferencias tales como las existentes en los lípidos, en las secuencias del ARN de transferencia y del ARN ribosómico, así como en los patrones de modificación y en la estructura de la subunidad enzimática.

Podría parecer que la naturaleza de la evolución, hace unos cuatro mil millones de años, era muy diferente de lo que fue más tarde. Los organismos sometidos a evolución serían entonces también muy distintos. Quizás el antepasado universal se encontrara en el proceso de desarrollar la pared celular, pues las paredes celulares archibacterianas difieren tanto de las verdaderas paredes bacterianas como las paredes celulares eucarióticas. Tal vez el antepasado universal estaba todavía desarrollando o refinando rutas bioquímicas; los lípidos se sintetizan de manera distinta en los dos reinos bacterianos, y muchos coenzimas son diferentes. Si llegara a descubrirse que las archibacterias difieren de las verdaderas bacterias en sus mecanismos para controlar la expresión de los genes (una posibilidad que aún ha de investigarse), podría entonces considerarse que el antepasado común disponía solamente de mecanismos rudimentarios de control genético.

La cuestión clave es si el antepasado universal estaba todavía desarrollando la unión genotipo-fenotipo cuando dio origen a sus líneas de descendencia. Dos observaciones sugieren que esto pudo haber sido así. La ARN polimerasa es un enzima que transcribe el gen a su complemento de ARN mensajero (que luego se traduce a proteína). La estructura de la subunidad de la ARN polimerasa es bastante constante entre las bacterias verdaderas, lo que no ocurre con la estructura de la polimerasa archibacteriana. ¿Podemos suponer que la función de la ARN polimerasa estaba todavía refinándose cuando se separaron las dos líneas bacterianas?

La segunda observación se refiere a los nucleótidos modificados de los

	ARCHIBACTERIAS	EUBACTERIAS	EUCARIOTAS
TAMAÑO CELULAR (DIMENSION LINEAL)	ALREDEDOR DE 1 MICROMETRO	ALREDEDOR DE 1 MICROMETRO	ALREDEDOR DE 10 MICROMETROS
ORGANULOS CELULARES	CARECEN	CARECEN	POSEEN
MEMBRANA CELULAR	CARECEN	CARECEN	POSEEN
PARED CELULAR	DIVERSOS TIPOS; NINGUNO TIENE ACIDO MURAMICO	VARIEDAD DENTRO DE UN TIPO; TODAS TIENEN ACIDO MURAMICO	NO HAY PARED CELULAR EN LAS CELULAS ANIMALES; DIVERSOS TIPOS EN OTROS PHYLA
LIPIDOS DE LA MEMBRANA	CADENAS ALIFATICAS UNIDAS POR ENLACES TIPO ETER	CADENAS ALIFATICAS RECTAS UNIDAS POR ENLACES ESTER	CADENAS ALIFATICAS RECTAS UNIDAS POR ENLACES ESTER
ARN DE TRANSFERENCIA: TIMINA EN BRAZO "COMUN"	FALTA	PRESENTE EN LA MAYORIA DE LOS ARNt DE LA MAYORIA DE LAS ESPECIES	PRESENTE EN LA MAYORIA DE LOS ARNt DE TODAS LAS ESPECIES
DIHIDROURACILO	FALTA, SALVO EN UN GENERO	PRESENTE EN LA MAYORIA DE LOS ARNt DE TODAS LAS ESPECIES	PRESENTE EN LA MAYORIA DE LOS ARNt DE TODAS LAS ESPECIES
AMINOACIDO DEL ARN DE TRANSFERENCIA INICIADOR	METIONINA	FORMILMETIONINA	METIONINA
RIBOSOMAS: TAMAÑOS DE LAS SUBUNIDADES	30*S*, 50*S*	30*S*, 50*S*	40*S*, 60*S*
LONGITUD APROXIMADA DEL ARN 16*S* (18*S*)	1500 NUCLEOTIDOS	1500 NUCLEOTIDOS	1800 NUCLEOTIDOS
LONGITUD APROXIMADA DEL ARN 23*S* (25-28*S*)	2900 NUCLEOTIDOS	2900 NUCLEOTIDOS	3500 NUCLEOTIDOS, O MAS
FACTOR DE TRADUCCION-ELONGACION	REACCIONA CON LA TOXINA DIFTERICA	NO REACCIONA CON LA TOXINA DIFTERICA	REACCIONA CON LA TOXINA DIFTERICA
SENSIBILIDAD AL CLORANFENICOL	INSENSIBLES	SENSIBLES	INSENSIBLES
SENSIBILIDAD A LA ANISOMICINA	SENSIBLES	INSENSIBLES	SENSIBLES
SENSIBILIDAD A LA KANAMICINA	INSENSIBLES	SENSIBLES	INSENSIBLES
SITIO DE UNION AL ARN, *AUCACCUCC* EN EL EXTREMO 3' DEL ARN 16*S* (18*S*)	POSEEN	POSEEN	CARECEN

RASGOS MOLECULARES de las archibacterias, las eubacterias y los eucariotas. Esas características permiten distinguir claramente los tres reinos primarios. En algunos aspectos, las archibacterias son únicas; en otros, se asemejan a las eubacterias o a los eucariotas (*color*). La timina y el dihidrouracilo son bases modificadas que sustituyen al uracilo en los ARN de transferencia. El factor de elongación es un componente de la maquinaria de traducción.

ARN de transferencia y ribosómicos. Tal como he indicado antes, los patrones de modificación son casi invariables dentro de cualquier reino primario, aunque tienden a diferir entre reinos. A pesar de que se desconoce la función de los nucleótidos modificados de los ARN de transferencia y ribosómicos, parece razonable presumir que la mayoría de ellos sirven para "afinar" la traducción, y hacerla así más precisa. De ser ello cierto, resultaría que muchas modificaciones han evolucionado independientemente en cada uno de los reinos primarios. La independencia de estas modificaciones implica, a su vez, que el antepasado universal no poseía las moléculas altamente especializadas del ARN de transferencia y ARN ribosómico de hoy, sino que disponía de una maquinaria de traducción más rudimentaria. Así las cosas, sólo podemos decir que los hechos están de acuerdo, y ciertamente sugieren, que el antepasado universal debió emplear mecanismos de transcripción y de traducción rudimentarios, esto es, que era un progenote.

El origen del urcariota

En base al tipo de ARN ribosómico de que disponen, las archibacterias, las eubacterias y los eucariotas parecen ser genealógicamente equidistantes entre sí; no se ha detectado ninguna relación entre dos cualesquiera de ellos. Sin embargo, en términos de la secuencia de aminoácidos de la proteína ribosómica *A*, las archibacterias parecen claramente estar emparentadas con las eucariotas. Pudiera ser, por tanto, que las archibacterias participasen, con las bacterias verdaderas, en la formación de la célula eucariota. Tal vez sea en las archibacterias donde debiéramos buscar el origen de un estadio de la célula eucariótica que queda todavía por explicar: el urcariota, que sirvió de huésped a los antepasados endosimbiontes de la mitocondria y del cloroplasto. (En realidad, más que buscar un huésped hipotético, habría que preguntarse si existió tal entidad. No es este el momento de dar forma a nuevos descubrimientos a partir de viejos prejuicios.)

Tal como ya he indicado, las diferencias entre la célula eucariota y los demás tipos principales de células, a nivel molecular, son más extensas y profundas que cualquiera de las diferencias apreciables al microscopio. El núcleo eucariótico parece contener, al menos, tres clases de genes: los de origen eubacteriano (seguramente procedentes de los genomas de los orgánulos de los eucariotas), los de origen archibacteriano (por ejemplo, el gen de la proteína ribosómica *A*) y los de un tercer origen, no identificado (ejemplarizados por el ARN ribosómico del citoplasma). ¿En qué medida es el núcleo eucariótico una quimera genética, esto es, una entidad compuesta por partes acopladas procedentes de fuentes distintas? ¿En qué estadio (o estadios) de la evolución tuvo lugar el presunto acoplamiento? ¿Y cuál era la natúraleza de los organismos que proporcionaron los diversos genes y estructuras?

Los biólogos se han inclinado a mirar a la célula eucariota como resultante de la asociación de células procariotas totalmente evolucionadas; se supone que su asociación creó un tipo de célula "superior", la célula eucariota. (El término procariota, "antes del núcleo", encierra estas implicaciones.) Quizá debamos plantearnos si los acontecimientos evolutivos que dieron a la célula eucariota sus características moleculares básicas eran realmente de esta naturaleza. La célula eucariota parece ser una quimera a un nivel muy básico. El propio ribosoma eucariota parece ser quimérico, al proceder su ARN de una fuente distinta de la de una, al menos, de sus proteínas. Si esta interpretación de los datos es correcta (y las futuras investigaciones lo han de aclarar) puede que la célula eucariota sea una entidad diferente de lo que hoy se cree. Quizá fue quimérica antes de que alcanzase un estado de complejidad comparable al de los procariotas actuales; pudo haberlo sido al surgir de la condición del progenote. Más que una avanzada forma "superior", la célula eucariota podría representar un retroceso en la dinámica evolutiva de su antepasado desaparecido hace mucho, el progenote.

De todos los descubrimientos recientes en filogenia molecular, quizá sea lo más atractivo la rica información sobre los estadios precoces de la evolución que encierra la célula. No nos hallamos limitados ya a la mera especulación para explicar los orígenes de la vida. Se repite hasta la saciedad que los últimos decenios de nuestro siglo conocerán una época de la biología en que "la ingeniería genética" registrará avances espectaculares en medicina e industria. Quizá valga la pena caer en la cuenta de que la biología se halla ya en el umbral de una revolución más silenciosa, una revolución que le permitirá al hombre descubrir y comprender las raíces de la vida y, así, adquirir un conocimiento más profundo del proceso evolutivo.

Asignación de recursos mediante programación lineal

Estructuras poliédricas abstractas en espacios de muchas dimensiones pueden ser útiles en planificación y gestión. Un algoritmo nuevo ha logrado acotar superiormente la complejidad de tales problemas

Robert G. Bland

Examinemos el caso de una pequeña fábrica de cervezas, cuyas bebidas tuvieran siempre demanda sobrada, pero cuya producción se encontrase limitada por escasez de ciertas materias primas. Supongamos que los ingredientes escasos sean maíz en grano, lúpulo y malta de cebada. Las proporciones de éstos dependen del tipo de cerveza a elaborar. Por ejemplo, para un barril de cerveza negra, del tipo llamado *ale* en Inglaterra, se emplea más malta que para las cervezas rubias, tipo *pilsen*. Además, el fabricante logra beneficios de 13 y 23 dólares por barril de *ale* y *pilsen*, respectivamente. Sujeto a estas condiciones, ¿cómo debe actuar el cervecero para lograr beneficios máximos?

Podría parecer que el cervecero debería dedicar todos sus recursos a la fabricación de *pilsen*, pues en ellas es máximo el beneficio por barril. Empero, tal decisión pudiera ser muy poco aconsejable, pues la elaboración de cerveza rubia pudiera consumir algunos de los recursos escasos disponibles mucho antes que la fabricación de *ale*. Si, por ejemplo, para producir un barril de *ale* se precisasen 5 kilos de maíz y para igual cantidad de cerveza rubia hiciesen falta 15, podríamos hacer tres veces más *ale* que cerveza. Además, si sólo preparase cerveza, el fabricante podría encontrarse con sus existencias de maíz agotadas mucho antes de que se le acabasen el lúpulo y la malta. Tal vez resulte que para sacar de sus existencias el mejor provecho posible deba dedicarlas en parte a cerveza y en parte a *ale*, para lograr así máximos beneficios. La determinación del programa de producción óptimo no es problema trivial. Y éste es justamente el tipo de problemas que puede ser resuelto gracias a las técnicas de programación lineal.

La programación lineal es una rama de las matemáticas que se ocupa de formular explícitamente y analizar cuestiones como la anterior. A su vez, forma parte de otra teoría más amplia, llamada investigación operativa, que se vale de diversos métodos de construcción de modelos matemáticos y de análisis cuantitativo para resolver problemas en organizaciones y grandes empresas. La programación lineal fue desarrollada poco después de la Segunda Guerra Mundial, en respuesta a los problemas logísticos que plantearon la guerra y los acontecimientos inmediatamente posteriores. Así, una de las primeras publicaciones sobre aplicaciones de la programación lineal fue un análisis del puente aéreo tendido cuando el bloqueo de Berlín, en 1948.

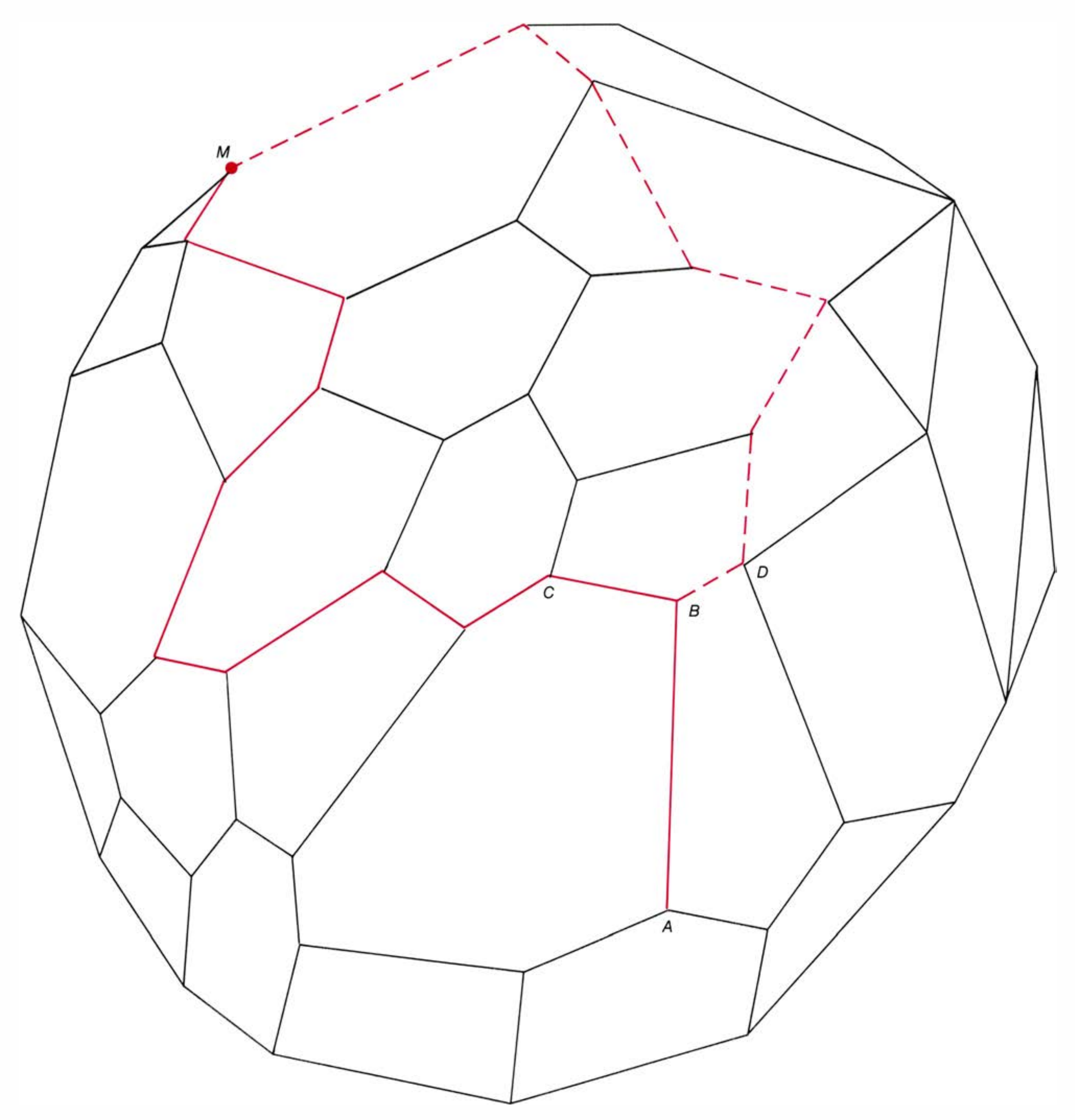
Si bien la resolución efectiva de los problemas de programación lineal exige casi siempre, como herramienta indispensable, el empleo de ordenador, el término "programación" se emplea aquí en el sentido de planificación, y no hace referencia a la programación de ordenadores. "Lineal", por otra parte, hace alusión a una propiedad matemática que presentan muchos problemas, y que facilita y simplifica su análisis. En el problema de la cervecería se presume que las cantidades de ingredientes o recursos necesarios para producir bien *ale*, bien cerveza, son proporcionales a las cantidades de bebida producidas. Duplicar la producción de cerveza exige duplicar las cantidades de ingredientes necesarios, y se supone que duplica también las ganancias logradas por su venta. Representando gráficamente la cantidad de maíz consumido en la preparación de cerveza en función de las cantidades producidas, la gráfica resulta rectilínea. Para poder aplicar las técnicas de programación lineal es necesario suponer, además, que tanto recursos como productos son divisibles, al menos aproximadamente. Por ejemplo, ha de poderse producir medio barril de cerveza, y su valor ha de ser mitad del valor del barril completo.

Por lo común, los problemas de programación lineal están ligados a problemas de asignación o reparto de bienes escasos entre cierto número de productos o actividades, con las hipótesis de divisibilidad y proporcionalidad descritas ya. Los recursos escasos pueden ser materias primas, productos semielaborados, mano de obra, capital a invertir o tiempos de procesado en máquinas grandes o costosas. Una distribución óptima puede ser la que haga máxima cierta medida del beneficio o utilidad –las ganancias, por ejemplo– o bien, la que haga mínima cierta medida de costo. En nuestros días, con productividad menguante y crisis de recursos financieros y materiales, tal vez valga la pena echar un vistazo a técnicas que ayuden a distribuir con máxima eficacia los recursos disponibles.

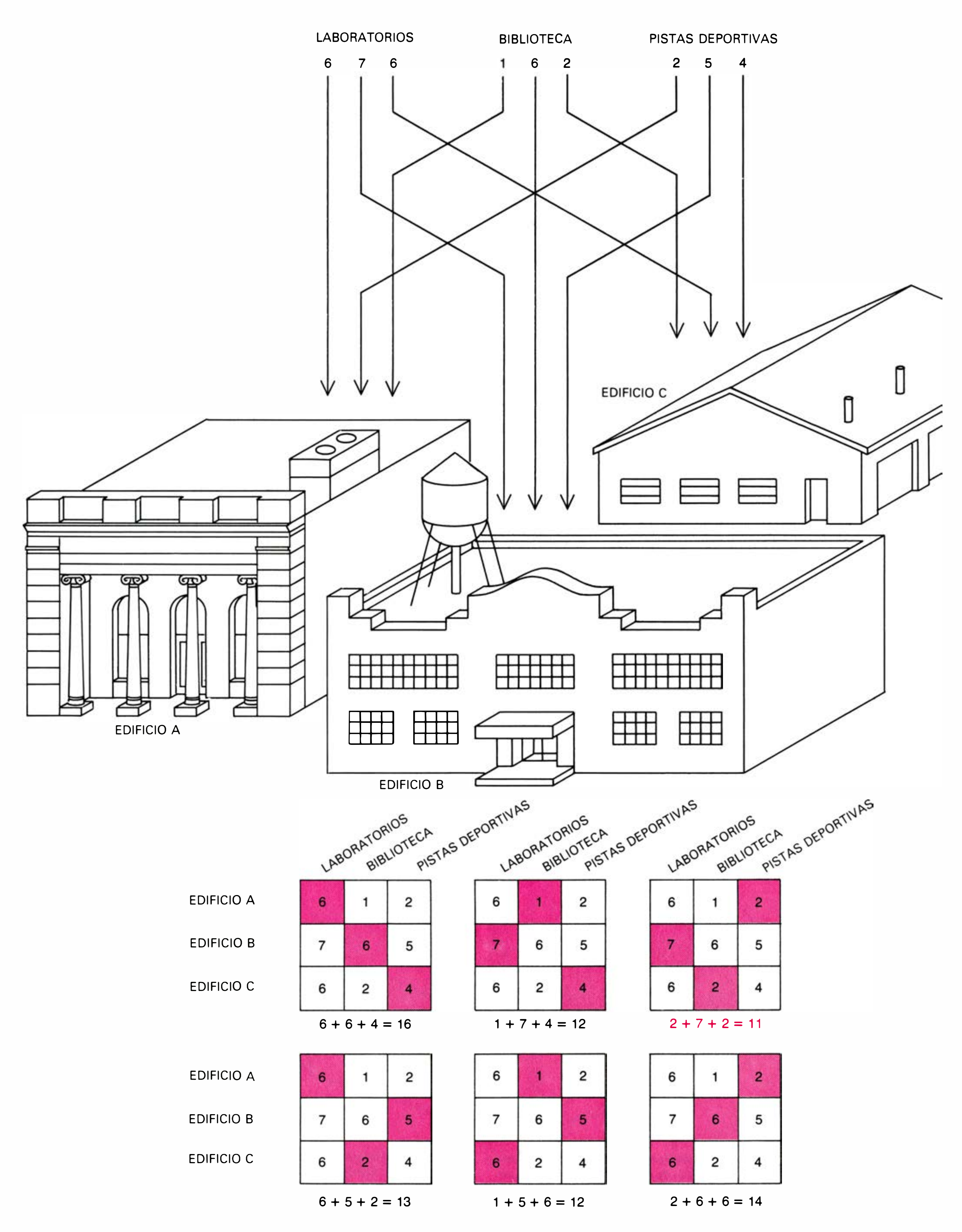
Las propiedades de los problemas de programación lineal se deducen de principios elementales de álgebra y geometría. Para resolver eficientemente estos problemas se cuenta con algoritmos –procedimientos detallados paso a paso– que aprovechan con gran ingenio estos principios algebraicos y geométricos. También los algoritmos son de concepción sencilla, aunque los detalles de su manejo pueden alcanzar niveles de gran complejidad. La importancia que reviste la programación lineal en economía es en gran medida consecuencia de la versatilidad y eficiencia de un solo algoritmo, el llamado "método del símplex".

El método del símplex (o símplice)

fue dado a conocer en 1947 por su creador, George B. Dantzig, actualmente en la Universidad de Stanford. Su fundamental importancia está justificada por su rapidez, por la riqueza de sus aplicaciones y porque es capaz de dar respuesta a cuestiones importantes, como por ejemplo la sensibilidad de las soluciones frente a variaciones de los datos de entrada. Tales cuestiones podrían adoptar la forma siguiente: ¿Cómo debería reajustarse el plan de producción del cervecero si cambiaran, por ejemplo, la disponibilidad de lúpulo o la ganancia por barril? ¿Cuánto debería estar dispuesto a pagar, como máximo, para lograr mayores suministros de los bienes escasos? ¿Qué precios debería pedir a otros empresarios interesados en adquirir parte de sus existencias? El método de símplex puede ser de ayuda a la hora de decidir si debemos comprar una máquina nueva o vender una de las que ya poseemos; si convendrá tomar un préstamo o prestar a terceros de nuestro capital; si contratar horas extraordinarias o suprimirlas. Con el mé-



METODO DEL SIMPLEX para resolver problemas de programación lineal. A fin de descubrir alguna de las asignaciones de recursos óptimas van recorriendo vértices de un politopo a lo largo de una línea quebrada cuyos lados son aristas. El politopo que aquí vemos es un poliedro tridimensional, y sus caras son poligonales. Cada uno de los puntos de la región donde se ha construido el politopo equivale a un plan –o programa– de asignación de mano de obra, capital u otros recursos. Cada uno de estos programas conlleva un beneficio o un costo netos. Con la programación lineal se pretende determinar programas de beneficio máximo o costo mínimo. El politopo define la región de factibilidad: el conjunto de todos los planes de distribución de recursos representados por puntos del interior o de la superficie del politopo son realizables, mientras los representados por puntos del exterior son imposibles, por insuficiencia de algún recurso. Cuando la relación entre recursos y beneficios (o costos) es lineal, los valores óptimos de éstos se encuentran necesariamente en vértices del politopo. El algoritmo del símplex va examinando los vértices conforme a una regla de selección, describiendo una trayectoria poligonal (*ABC...M*) por las aristas del politopo. Cada uno de los vértices de esta ruta va mejorando los beneficios o costos del anterior, hasta alcanzar un punto *M* donde se logra el máximo o el mínimo. Hay muchos caminos que partiendo de *A* conduzcan a *M*. El politopo puede ser pluridimensional.



ESTE PROBLEMA DE ASIGNACION busca hacer mínimos los costes de adaptación de tres edificios a tres funciones (laboratorio, biblioteca y pista cubierta), estando cada edificio dedicado tan sólo a una de ellas. En el caso de tres edificios y tres funciones es preciso tomar en consideración $3^2 = 9$ costos de reconversión. Estos costos (en millones de dólares) pueden ordenarse cómodamente en una matriz; cualquier asignación factible consistirá en tomar de esta matriz un número de cada fila y uno de cada columna, lo que puede hacerse de $3 \times 2 \times 1 = 6$ formas. En general, para problemas de asignación de tamaño n por n, el número de posibles asignaciones es "factorial de n" (escrito $n!$) que es igual al producto de n por todos los enteros positivos menores que n. Dado que el valor de $n!$ crece muy rápidamente, no es práctico determinar el costo mínimo por enumeración de todas las asignaciones.

todo del símplex podemos introducir en el problema restricciones adicionales, y resolver nuevamente el problema para apreciar sus efectos. Por ejemplo, el método puede decirle rápidamente a un empresario cuál será el costo de proporcionar servicios poco rentables con la finalidad de conservar la buena disposición de un cliente.

El método del símplex ha dado prueba de su extraordinaria eficiencia para resolver complejos problemas de programación lineal con miles de restricciones. A los teóricos, empero, la velocidad con que resuelve tales problemas les resulta un tanto desconcertante. No disponemos de explicación definitiva de por qué es así. En realidad, los matemáticos han ideado problemas donde el método del símplex revela una lentitud intolerable. Al parecer, por razones que no están claras, tales problemas no se presentan en la práctica.

Recientemente, en Rusia se ha preparado un nuevo algoritmo de programación lineal, que en cierto sentido logra eludir algunas de las dificultades teóricas atribuidas al método de Dantzig. Tal logro fue publicado en las primeras planas de muchos periódicos de todo el mundo, lo que da idea de la importancia que hoy se concede en economía a la programación lineal. Desdichadamente, el nuevo algoritmo, llamado "método del elipsoide", no ha dado hasta el momento pruebas de mejor eficiencia que el método del símplex a la hora de resolver problemas prácticos. Hay por ahora una curiosa discrepancia entre las medidas práctica y teórica del rendimiento de cómputo de los algoritmos.

Incluso empleando algoritmos eficientes, los costos de planteamiento de un problema de programación lineal de gran tamaño pueden ser considerables. Expresar mediante programación lineal un conjunto de circunstancias prácticas no es empresa sin escollos, ni tampoco lo es la recogida y organización de los datos que expresan tales circunstancias. Además, ya por fin en enunciado matemático, para resolver el problema se precisa un ordenador de gran potencia.

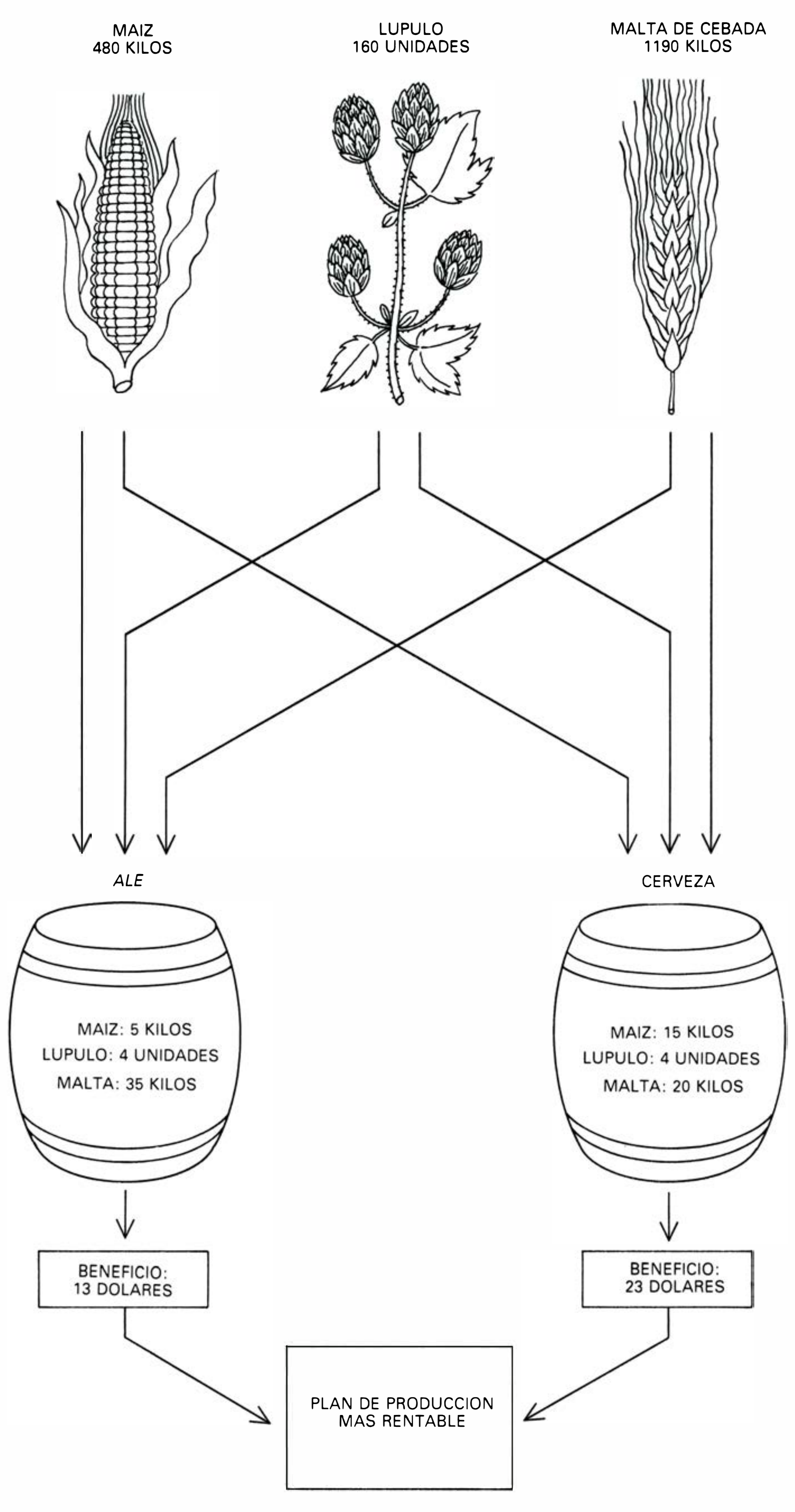
No obstante estos inconvenientes, los beneficios implicados por la programación lineal suelen ser importantes. A mediados de los años 50, cuando se pusieron a punto métodos de programación lineal para optimizar el mezclado de gasolinas, la Exxon Corporation empezó a economizar de un 2 a un 3 por ciento en las operaciones de mezclado. En la industria petrolífera, este tipo de aplicaciones tuvieron rápida difusión, y pronto se utilizaron para control de otros procesos, entre ellos, los de "cracking" catalítico, destilación y polimerización. Aproximadamente al mismo tiempo, otras industrias, sobre todo las papeleras y las cadenas de alimentación, así como las empresas siderometalúrgicas y las explotaciones agrícolas, empezaron a implantar métodos de programación lineal. Según Charles Boudrye, de Linear Programming, Inc., con sede en Silver Springs, Maryland, un determinado fabricante de papel logró en sólo un año aumentar en 15 millones de dólares sus beneficios, valiéndose de programación lineal para organizar y racionalizar su gama de productos.

Hay actualmente alrededor de una decena de compañías que ofrecen comercialmente programas de ordenador "a la medida" (que en el argot informá-

	RESTAR 3	RESTAR 4	RESTAR 8	RESTAR 2	RESTAR 9	RESTAR 7	RESTAR 0	RESTAR 5	RESTAR 1	RESTAR 6
RESTAR 2	6	6	12	4	14	11	3	9	4	10
RESTAR 5	8	12	17	8	16	16	6	13	7	13
RESTAR 6	11	13	18	9	15	16	8	14	8	14
RESTAR 4	8	10	15	10	16	13	7	9	7	11
RESTAR 8	13	15	16	11	19	19	10	16	12	18
RESTAR 5	9	12	17	8	16	15	7	13	7	11
RESTAR 3	8	8	13	8	13	12	6	8	4	11
RESTAR 1	6	8	11	3	14	10	2	8	3	9
RESTAR 0	5	5	12	2	12	7	2	9	6	12
RESTAR 7	15	14	17	13	17	18	7	18	10	16

1	0	2	0	3	2	1	2	1	2
0	3	4	1	2	4	1	3	1	2
2	3	4	1	0	3	2	3	1	2
1	2	3	4	3	2	3	0	2	1
2	3	0	1	2	4	2	3	3	4
1	3	4	1	2	3	2	3	1	0
2	1	2	3	1	2	3	0	0	2
2	3	2	0	4	2	1	2	1	2
2	1	4	0	3	0	2	4	5	6
5	3	2	4	1	4	0	6	2	3

PARA CONVENCER a un escéptico de que determinada distribución es óptima (*casillas coloreadas en la figura superior*) no se precisa evaluar el costo de todas las distribuciones posibles. En este ejemplo, el tamaño de la matriz de costes es de 10 por 10, y hay por tanto 10! = 3.628.800 posibles distribuciones. Al restar un mismo número a todos los elementos de una fila o columna, el costo total de cualquiera de las posibles asignaciones queda rebajado en igual valor, por lo que la ordenación relativa de los costos totales se conserva. Ello se debe a que en cada asignación se toma exactamente un número de la fila o columna modificada en la matriz. El conjunto de números a restar puede elegirse de manera que la matriz primitiva quede transformada en otra sin elementos negativos donde, además, cada fila y cada columna contenga el menos una casilla de coste nulo (*figura inferior*). Puesto que ninguna asignación puede tener coste total menor que cero, la asignación óptima correspondiente a la matriz de costes transformada habrá de tener coste nulo. Se deduce de aquí que también eran óptimos los cuadros señalados en color en la matriz primitiva. Para generar el sistema de números a sustraer se han ideado algoritmos eficientes.



EL DILEMA DEL CERVECERO permite explicar cómo se aplica la programación lineal para optimizar la distribución de recursos entre varios procesos de fabricación. El cervecero ve limitada su producción de cerveza y *ale* por la escasez de tres ingredientes esenciales: maíz, lúpulo y malta de cebada. Los niveles de producción factibles quedan determinados no sólo por las existencias de cada ingrediente, sino también por las proporciones en que éstos entran en cada producto elaborado. La función objetivo, es decir, la magnitud que debe hacerse óptima, es el beneficio. En programación lineal se supone que todos los recursos, los productos y las ganancias son indefinidamente divisibles: el cervecero puede consumir medio kilo de maíz, vender un cuarto de barril de *ale* y materializar la parte proporcional de beneficios.

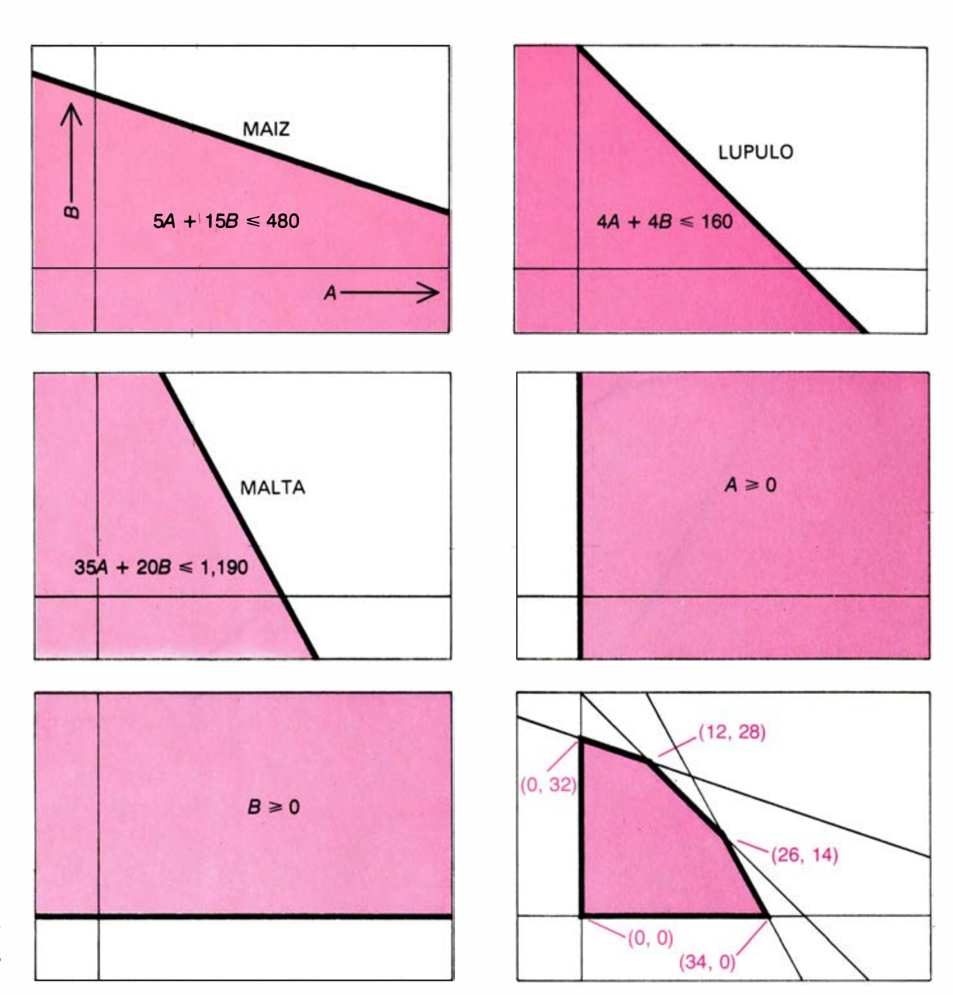
tico a veces llaman "packages"), basados en el método del símplex. Tales programas prestan servicio directo a unos 1000 clientes, quienes los utilizan con licencia de los programadores. Estos clientes abonan sustanciosas cuotas mensuales, por lo que es de suponer que se valdrán del método con asiduidad. Otras muchas organizaciones y empresas tienen acceso a estos "packages" por intermedio de entidades de gestión y asesoramiento. Además, se han puesto a punto diversos programas especializados para análisis de flujo a través de redes, programas que seguramente prestan servicios todavía mayores que los algoritmos de uso general que hemos mencionado.

Puede bastar una sola organización para poner de relieve la variedad y amplitud de las aplicaciones de la programación lineal. Así, en la actualidad, Exxon Corporation se vale de ella para planificar sus prospecciones y sondeos; para la distribución de crudos a las distintas refinerías; para determinar las condiciones de proceso en cada una; en la distribución de los productos y en la planificación misma de su estrategia comercial. Según reseña David Smith, del departamento de ciencias de cómputo y comunicación de Exxon, la programación lineal y sus ramificaciones consumen de un 5 a un 10 por ciento de las tareas de cómputo de la compañía. Tal proporción se ha mantenido a lo largo de los últimos 20 años, a pesar del rápido desarrollo que en todos los órdenes ha experimentado el procesamiento automático de información.

Aunque el método del símplex sea instrumento de gran potencia y versatilidad, está fundado en ideas de carácter elemental. Para mejor comprender algunas de ellas convendrá examinar un problema de estructura especial, que suele denominarse "problema de asignación". Fijémonos en el ejemplo siguiente: el patronato de una universidad tiene tres edificios pendientes de remodelación, y tres necesidades que cubrir, para las cuales se propone utilizar los edificios. Supongamos que las necesidades sean laboratorios, biblioteca y pistas deportivas cubiertas. Cada edificio puede solamente desempeñar una función. En la ilustración de la página 64 se han tabulado los costos de transformación asociados con cada uno de los nueve emparejamientos posibles de edificio y finalidad del mismo. ¿De qué forma podrá el patronato gestionar la remodelación de los pabellones, a costo mínimo?

El problema puede resolverse señalando en el casillero tres de los nueve cuadros que contiene. Hay que elegir exactamente uno en cada fila y en cada columna, de forma que cada local esté dedicado a un determinado servicio, y cada función tenga acomodo en uno de los edificios. La solución más económica –óptima– será aquélla donde la suma de costos anotados en las casillas sea mínima. En este caso, no es difícil hallar la solución económicamente óptima, porque sólo hay unas cuantas posibilidades en la elección de las casillas. Una vez señalado uno de los cuadros de la primera fila, tan sólo quedan dos para elegir en la segunda; en la tercera, la elección ya es obligada, porque tan sólo queda en ella un cuadrado no situado en una de las dos columnas ya seleccionadas. Hay por consiguiente 3×2×1, es decir, 6 formas de efectuar la distribución de funciones, y es cosa fácil calcular el costo total correspondiente a cada una, y seleccionar la menos costosa.

Si bien este análisis, por enumeración exhaustiva de los casos posibles, resuelve rápidamente el problema cuando la matriz sea de 3 por 3, no resultará práctico en problemas de dimensión mayor. Supongamos que los edificios y funciones fueran cuatro. Ahora el número de asignaciones posibles de funciones a edificios es de 4×3×2×1, o sea, 24. Enunciado en general, el problema es asignar n funciones a n edificios, y el número de posibles distribuciones es *factorial de n* (denotado $n!$), igual al producto de n por todos los enteros anteriores, desde 1 hasta $n - 1$. Cuando n es 10, hay 10! asignaciones distintas, o sea, más de 3,6 millones de posibilidades diferentes.



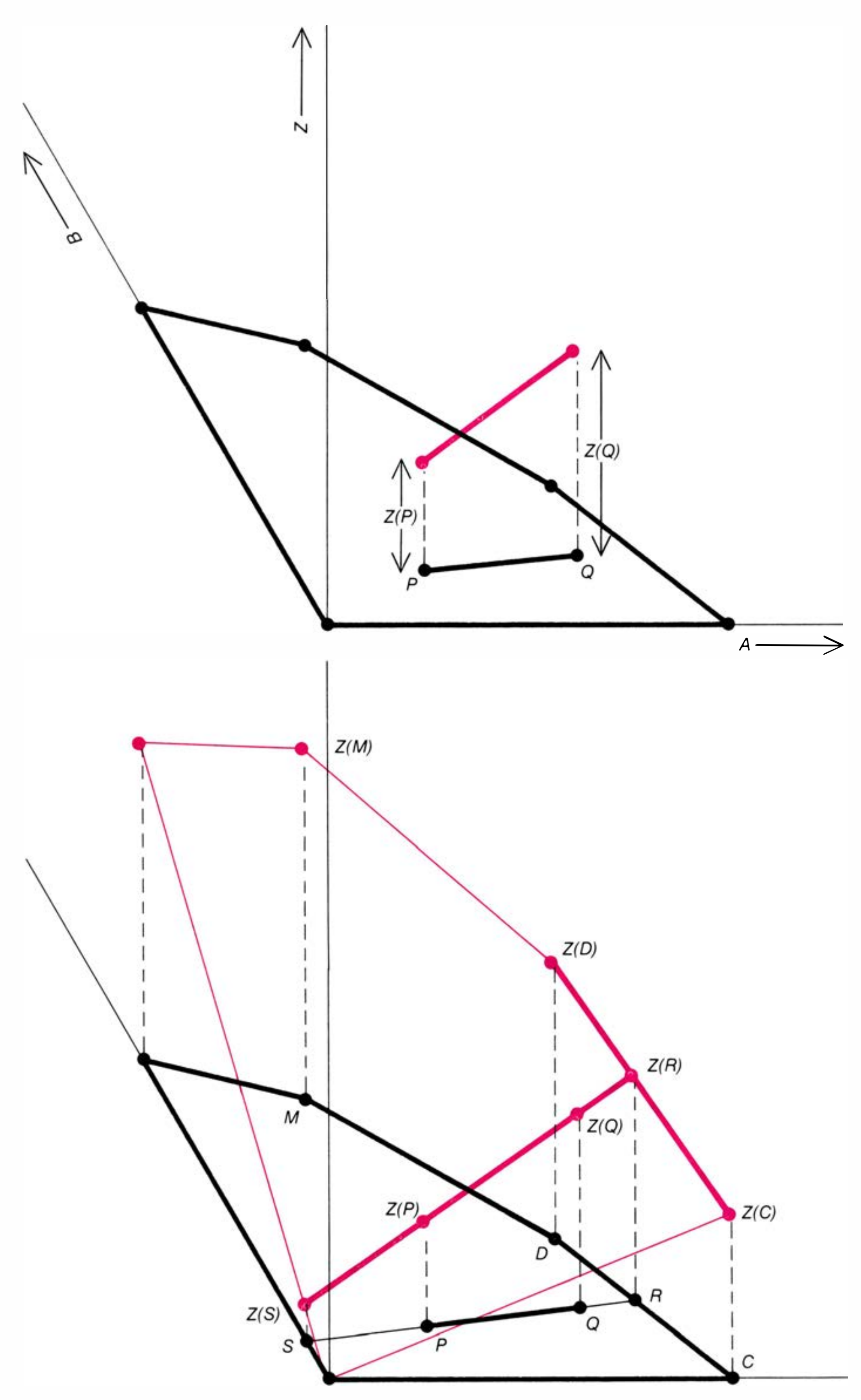
LA REGION FACTIBLE asociada al problema del cervecero es intersección de cinco semiplanos. Cada punto (A,B) del plano equivale a un plan de producción consistente en elaborar A barriles de *ale* y B barriles de cerveza. Los tres primeros semiplanos representan gráficamente todos los programas de producción realizables, dadas las existencias de cada ingrediente. Por ejemplo, la cantidad de maíz necesaria para el programa (A,B) es $5A + 15B$ (peso en kilos del maíz preciso para un barril de *ale* por el número de barriles A, más el peso por barril de cerveza por su número B de barriles). Esta cantidad no puede superar el total de 480 kilos, que son las existencias del cervecero. Por consiguiente, cualquiera de los puntos situados por debajo de la recta $5A + 15B = 480$ representa un programa de producción que no requiere más maíz del disponible. Los semiplanos correspondientes al lúpulo y la malta se construyen igual. Los otros dos semiplanos expresan que sólo caben programas con producciones no negativas.

La $n!$ crece a tal velocidad que pronto convierte en desencanto cualquier posible entusiasmo que pudiera haber suscitado en nosotros el método de enumeración. Supongamos que tuviéramos que resolver un problema de asignación de dimensiones 35 por 35, que deseáramos hacerlo por el método de enumeración, y que para ello dispusiéramos de un ordenador capaz de ir examinando cada una de las distribuciones posibles, de evaluar sus costos y de irlos comparando con la distribución de costo mínimo descubierta hasta el momento. Supongamos que nuestro ordenador ejecutase esta tarea a razón de mil millones de casos por segundo. (Un ordenador capaz de semejante velocidad sería, con mucho, más rápido que ninguno de los actualmente disponibles.) Pues bien, incluso si la tarea de examinar las 35! asignaciones fuese encomendada simultáneamente a mil millones de ordenadores como el nuestro, tras mil millones de años de funcionamiento continuo tan sólo se habría examinado una porción insignificante del total de casos.

No se crea que un problema de asignación de 35 por 35 es un problema grande. Si, por ejemplo, la tarea consistiera en asignar puesto de trabajo a 35 personas de la plantilla de una empresa, de forma que los costos de formación y capacitación fuesen mínimos, es perfectamente posible el valor $n = 35$. Hay muchos otros problemas de asignación donde n es mayor o igual que 1000. Como es obvio, tales problemas exigen análisis un poco más penetrantes que la simple y pura enumeración.

La pesada carga de enumerar exhaustivamente todos los casos podría quedar muy aligerada si pudiéramos ahorrarnos el examen de aquellas asignaciones que resultaran más costosas que las ya estudiadas. Lograríamos tal efecto si dispusiéramos de una regla de detención del proceso, de un criterio de calidad o bondad de las asignaciones, que nos permitiera reconocer rápidamente que hemos acertado con una asignación óptima. Todo algoritmo que contenga criterios de este tipo conlleva ventajas secundarias de gran importancia. Tales ventajas quedan resumidas en lo que Jack Edmonds, de la Universidad de Waterloo en Ontario, ha denominado "el principio de supervisión absoluta", o lo que podríamos llamar también, "el problema del jefe incrédulo".

Supongamos que tras tediosa enume-



VALOR DE LA FUNCION OBJETIVO, representado gráficamente como una altura $Z(P)$ por encima o debajo de P, medida sobre el eje z. (Las distancias sobre el eje z no están dibujadas a la misma escala que las del plano de la región factible.) Podemos también imaginar que $Z(P)$ es un punto del espacio tridimensional situado sobre P a la altura ya explicada. Cuando la función objetivo es lineal, la porción de gráfica correspondiente a cualquier segmento de recta contenido en la región factible es también un segmento (*en color, en la figura superior*). Todo punto P interior a la región factible está contenido en una recta que intercepta en dos puntos a la frontera de la región (*puntos R y S de la ilustración inferior*). Si el segmento que une a través del espacio los puntos $Z(R)$ y $Z(S)$ no es paralelo al plano de la región factible, la función objetivo tendrá que tomar sus valores máximo y mínimo correspondientes al segmento en uno de los extremos, $Z(S)$ por ejemplo, asociado con el punto S del contorno de la región factible. También a lo largo de cada lado del contorno es rectilínea la gráfica de la función objetivo; por cuyo motivo, ésta alcanzará su máximo valor en alguno de los extremos, $Z(D)$ por ejemplo, que corresponde a uno de los vértices de la región factible. Así pues, el valor de la función objetivo en un punto interior de esta región es siempre superado en algún punto de uno de los lados, mientras que los valores en puntos de los lados se ven superados en los vértices de la región factible. Para hallar el valor máximo $Z(M)$ de la función objetivo es necesario examinar tan sólo los valores en los vértices. Cuando la región factible sea bidimensional, la gráfica de la función objetivo será un plano, cuya altura máxima se alcanzará sobre un vértice.

ración hemos conseguido resolver el problema de asignación de tamaño 10 por 10 que vemos en la ilustración superior de la página 65, examinando uno por uno la totalidad de los 3.628.800 posibles repartos. La asignación óptima, mantenemos nosotros, corresponde a la sombreada en color en la ilustración. Al presentarle la solución al jefe, éste nos escruta con desconfianza, da una chupada a su cigarro y nos pregunta "¿cómo sé yo que no hay otras soluciones menos costosas?". Y entonces, ya podemos ir tragando saliva, pues, según parece, la única forma de dejar manifiestos los méritos de nuestra solución sería repetir el examen de las 3.628.800 posibilidades, esta vez, bajo la mirada escrutadora del mandamás.

En cambio, con una regla de detención dispondríamos de una demostración breve de la imposibilidad de mejorar la solución. Supongamos que pudiéramos presentarnos ante el jefe no sólo con la asignación óptima, sino pertrechados, además, de un conjunto de números que restar de los elementos de filas y columnas. Para explicar cómo se obtienen estos números sería preciso analizar con detalle el problema de asignación; baste señalar que tal sistema de números puede ser determinado por ordenador, mediante un algoritmo eficiente. Una vez hallados, la utilidad de estos números se aprecia sin dificultad. Notemos, para empezar, que restándole un mismo número a todos los elementos de una fila o columna dada lo que hacemos es deducir esta cantidad del costo total de cada una de las asignaciones posibles. En efecto, en toda asignación factible es forzoso tomar en la matriz un elemento, y sólo uno, de cada fila y cada columna. Por ejemplo, al restarle 5 a todos los elementos de la sexta fila, cada una de las nuevas asignaciones posibles contendrá un elemento que es exactamente cinco unidades menor que el de la asignación correspondiente con la tabla de costos primitiva. Por consiguiente, la situación relativa de los costos de las diferentes asignaciones permanecerá invariable. Tal substracción puede efectuarse reiteradamente, con tal de que se aplique por igual a todos los elementos de una misma fila o columna.

Mediante sustracción reiterada podemos transformar la matriz de costos primitiva en la que vemos en la ilustración inferior de la página 65. Esta última tiene una propiedad muy notable, pues ahora podemos hacerle observar a nuestro jefe que los costos correspon-

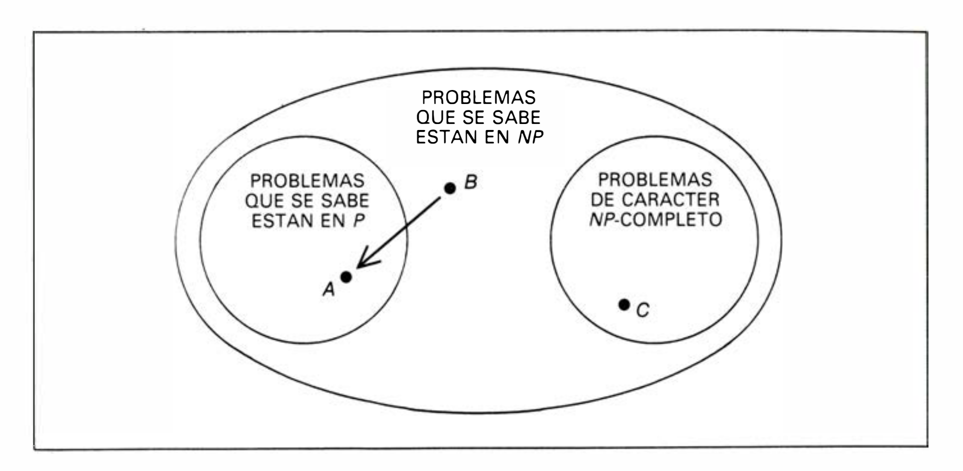
dientes a la distribución óptima sontodos cero, y que ninguno de los números de la matriz es negativo. Puesto que la suma de costos señalados en nuestra distribución es nula, y puesto que no hay costos negativos, ninguna de las asignaciones posibles tendrá coste inferior. En breve, basta hacer unos pocos cientos de cálculos –y no decenas de millones– para demostrar al supervisor más escéptico que ninguna asignación podrá valer menos que la elegida.

A pesar de no haber explicado con detalle cómo se resuelve un problema de asignación ni cómo se hallan los números a substraer de filas o columnas, el problema de asignación sí pone de relieve la necesidad de rehuir la enumeración exhaustiva, y también la posibilidad de lograrlo con ayuda de una regla de detención capaz de detectar el hallazgo de una solución óptima. Al diseñar algoritmos deben tomarse en cuenta estas consideraciones, no sólo al resolver el problema de asignación, sino, más generalmente, al aplicarlo a problemas de programación lineal.

Volvamos a fijarnos en el cervecero, cuyos dos productos, cerveza y *ale*, requieren en su elaboración distintas proporciones de maíz, lúpulo y malta. Supongamos que sus existencias actuales sean de 480 kilos de maíz, 160 unidades de lúpulo y 1190 kilos de malta; supongamos también que su producción esté limitada por la escasez de estas materias primas. Durante el proceso de elaboración se consumirán también otros recursos, como agua, levadura, energía y mano de obra; pero se supone que estos recursos son abundantes. Si bien sus costos pueden pesar en el ánimo del cervecero, no son causa limitante directa de su capacidad de producción. Supongamos que cada barril de *ale* requiera cinco kilos de maíz, cuatro unidades de lúpulo y 20 kilos de malta. Supongamos, además, que toda la cerveza y todo el *ale* que puedan fabricarse tengan salida a los precios habituales del mercado, que dejan respectivamente 13 y 23 dólares de beneficio por barril de *ale* y de cerveza.

La escasez de maíz, lúpulo y malta limita forzosamente los niveles de producción potencial. Por ejemplo, aunque se dispone de lúpulo y malta suficientes para preparar más de 32 barriles de cerveza, la elaboración de tan gran producción de este brebaje agotaría las existencias de maíz, impidiendo así la producción tanto de cerveza como de *ale*. Otro programa de fabricación viable sería no elaborar nada de cerveza y fabricar 34 barriles de *ale*, lo que dejaría exhaustas las existencias –1190 kilos– de malta. De ambas, la primera alternativa parece preferible. Con la primera línea de producción los beneficios serían de 32 × 23 = 736 dólares, mientras que la segunda posibilidad sólo daría 34 × 13 = 442 dólares.

Ahora, hay otros programas de producción, mixtos, más rentables que cualquiera de éstos. La fabricación de seis barriles de *ale* y 30 de cerveza consume por completo los 480 kilos de maíz, 154 de las 160 unidades de lúpulo y 810 de los 1190 kilos de malta, siendo los beneficios de (6 × 13 dólares) + (30 × 23 dólares), es decir, 768 dólares. Hay muchos programas capaces de beneficios todavía mayores. En este caso no sólo es poco práctico enumerar exhaustivamente todas las posibilidades; resulta, sencillamente, imposible. Existen infinitos programas compatibles con las condiciones del problema del cervecero. Cada uno de tales programas es una solución factible. Afortunadamente, hay un pequeño número de soluciones factibles especiales, llamadas soluciones extremales, que son las únicas que habrán de merecer nuestra atención.



LA TEORIA DE COMPLEJIDAD computacional ha conseguido recientemente situar la programación lineal entre el conjunto *P* de clases de problemas que admiten acotación polinómica. Tal logro (representado en la figura por la fecha que termina en *A*) ha sido posible gracias a una demostración del matemático ruso L. G. Khachian, quien ha probado que un nuevo algoritmo de programación lineal recientemente descubierto, llamado método del elipsoide, está acotado polinómicamente. Una función polinómica de la variable *n* es una suma finita de productos de números por potencias de *n*. Se dice que una clase de problemas está polinómicamente acotada cuando el número de operaciones aritméticas elementales necesarias para resolverlos está acotado por una expresión que es función polinómica de cierta medida *t* del tamaño del problema. Antes del resultado de Khachian se ignoraba si para la programación lineal podría existir algún algoritmo con esta propiedad. Se sabía ya que la programación lineal pertenecía al conjunto *NP* mucho más amplio de funciones no determinísticamente polinómicas (punto *B*). Hablando sin mucha precisión, *NP* es el conjunto de clases de problemas para los que puede comprobarse rápidamente si una solución propuesta es factible. Un segundo subconjunto de *NP* es el conjunto de problemas *NP*-completos; de lograrse una demostración de que una de las clases de problemas *NP*-completos está polinómicamente acotada, quedaría demostrado también que todas las clases de problemas *NP* se encuentran realmente en *P*. Los problemas *NP*-completos son, en cierto sentido, los problemas *NP* más difíciles (punto *C*).

La importancia de las soluciones extremales queda manifiesta al representar gráficamente, mediante puntos del plano, el conjunto de todas las soluciones factibles. La figura plana resultante se denomina región factible. Sea A el número de barriles de *ale* a elaborar según determinado plan de producción, y sea B el correspondiente número de barriles de cerveza. En programación lineal, las magnitudes de A y B se denominan variables de decisión; podemos asociarlas con cantidades tomadas sobre unos ejes de coordenadas en el plano. Todo punto del plano puede quedar determinado mediante un par ordenado (A,B) de números reales, que corresponderían también a un determinado sistema de niveles de producción.

Puesto que no serán posibles en este problema niveles de producción negativos, la región factible queda confinada en el primer cuadrante del plano, donde tanto A como B son no negativos. ¿Cómo afecta a la produción la escasez de malta? Sabemos que cada barril de *ale* precisa 35 kilos de malta, y que cada barril de cerveza requiere 20. La cantidad de malta necesaria para A barriles de *ale* y B de cerveza es, por tanto, $35A + 20B$. Si se van a consumir los 1190 kilos de malta disponibles, $35A \times 20B = 1190$. El conjunto de puntos (A,B) que satisfacen esta ecuación forma una línea recta. Todos los planes de producción (A,B) que exijan emplear más de 1190 kilos de malta co-

rresponden a puntos de uno de los semiplanos definidos por ella; los que requieran menos pertenecen al otro. A causa de la escasez de las existencias de malta, tan sólo estos últimos se hallan en la región factible, que contiene además los puntos situados en la propia recta.

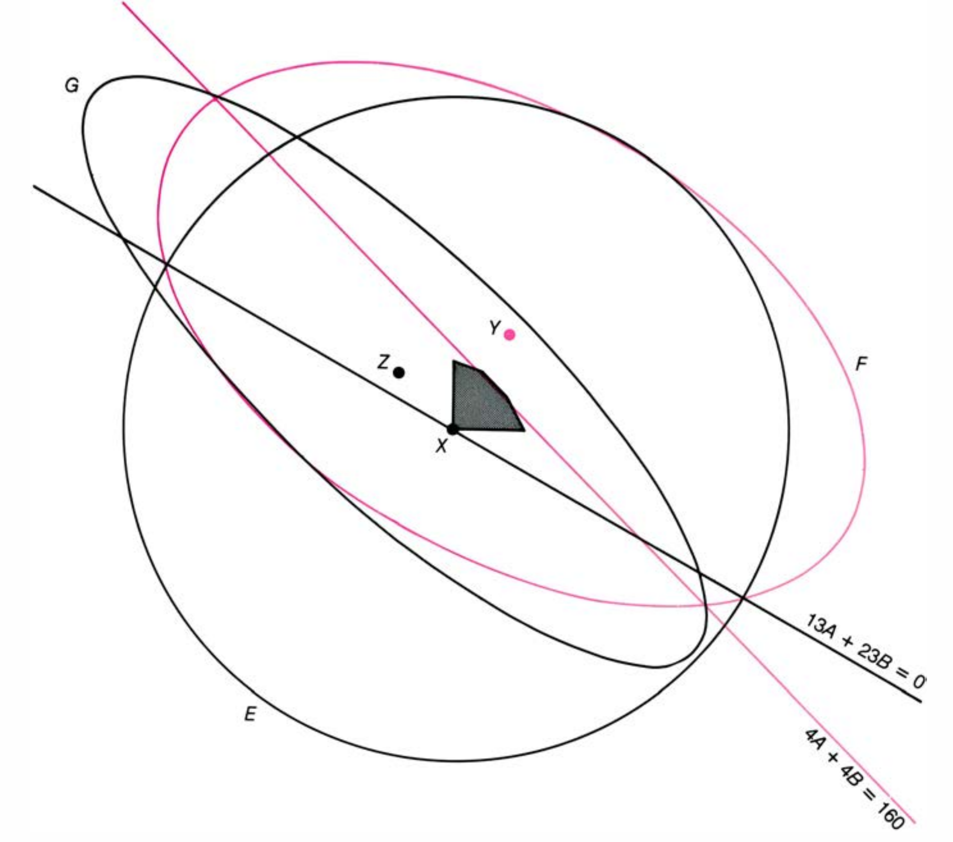
De igual manera, la escasez de lúpulo confina la región factible a uno de los semiplanos definidos por la recta $4A + 4B = 160$, y, la de maíz, a uno de los definidos por $5A + 15B = 480$. Aquellos puntos que satisfagan simultáneamente estas tres condiciones constituirán la región factible [*véase la ilustración de la página 67*]. Observemos que la región factible es convexa; es decir, todo segmento rectilíneo que conecte dos puntos de la región (aunque éstos se encuentren sobre los lados del polígono que la delimita) está enteramente contenido en ella.

Puesto que el cervecero gana 13 dólares por barril de *ale* y 23 en cada barril de cerveza, su problema es hacer máxima la ganancia conjunta por ambos conceptos: $13A + 23B$. Para conseguirlo tiene que hallar un punto (A,B) situado en la región factible donde la expresión $13A + 23B$ alcance un valor máximo. En programación lineal, esta función –los beneficios a hacer máximos, o los costes, a hacer mínimos– se denomina función objetivo.

Podemos incorporar la función objetivo a la representación gráfica de la región factible, añadiendo una tercera dimensión. Para cada punto (A,B) representante de un plan de producción, la ganancia esperada vendrá dada por la altura de la función $13A + 23B$ sobre el plano que contiene al punto. Evidentemente, la meta del cervecero es hallar un punto de la región factible donde la función objetivo alcance altura máxima. Si para ello tuviéramos que hallar el valor de la función en todos los puntos de la región, la tarea sería imposible; dos notables propiedades del problema sirven para delimitar muchísimo la zona a explorar. Estas propiedades son la convexidad de la región factible y la linealidad de la función objetivo.

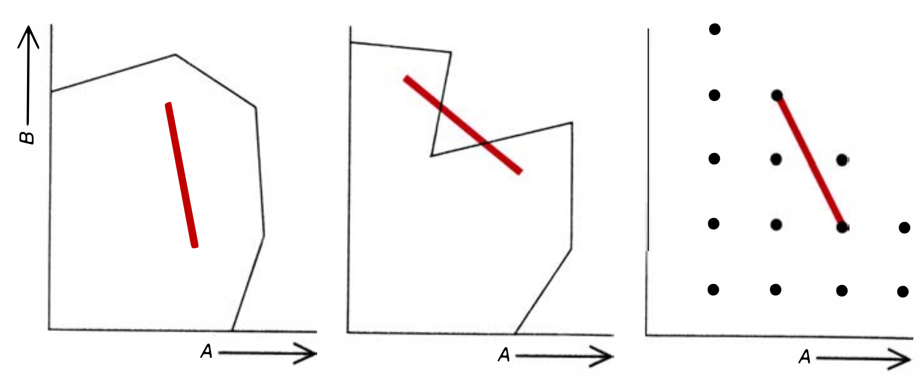
Por ser convexa la región, todo punto de su interior puede quedar incluido en un segmento rectilíneo cuyos extremos yacen en la frontera de la región. (En realidad, por cada punto interior pasa una infinidad de tales segmentos; carece de importancia cuál sea elegido.) Podemos construir la porción de gráfica de la función objetivo correspondiente al segmento, que será una línea en el espacio. Como la función objetivo es lineal, tal porción de gráfica es rectilínea [*véase la ilustración de la página 68*]. La gráfica puede resultar paralela al plano base, en cuyo caso todos los planes de producción correspondientes al segmento trazado en la región factible proporcionan iguales beneficios. Pero si la porción de gráfica de objetivo no fuese paralela al plano, su altura máxima se alcanzaría sobre uno de los extremos del segmento que, como ya sabemos, yace en la frontera de la región factible. Por tanto, puesto que el análisis anterior puede repetirse para todo segmento contenido en la región factible, resulta que el máximo absoluto de la función objetivo se hallará, inexorablemente, sobre la frontera de tal región. El cervecero, al buscar la máxima ganancia, puede despreciar todo el interior de la región factible y estudiar únicamente los planes de producción correspondientes a puntos situados en la frontera.

Pero el mismo razonamiento permite llevar el análisis un paso más allá. Al ser poligonal la región factible, cada punto de su frontera se encuentra en un segmento rectilíneo, cuyos extremos son sendos vértices del polígono. La porción de gráfica de la función objetivo correspondiente a uno de estos lados del polígono se construye de igual forma que para los segmentos que atraviesan el interior. Como antes, el máximo volverá a encontrarse en uno de los extremos (salvo que la función objetivo fuese constante sobre dicho lado). Por consiguiente, el valor máximo de la función objetivo ha de corresponder a alguno(s) de los vértices de la región factible. El cervecero tan sólo precisa, pues, calcular los valores de sus ganancias en, a lo sumo, todos los vértices del polígono, y seleccionar el que proporcione mayores beneficios. Podrá entonces estar seguro de que ningún otro



METODO DEL ELIPSOIDE, que, al igual que el método del símplex, puede ser interpretado geométricamente. En el problema del cervecero, comenzaría tomando un gran círculo *E* con centro en el origen, *X*, que contenga la región factible. El tamaño del círculo está determinado por los datos del problema. El método procede entonces a construir una sucesión de elipses cada vez menores, que contienen todas la solución óptima. Como el centro del círculo es factible, la siguiente elipse de la construcción habrá de contener todos aquellos puntos del círculo donde la función objetivo sea al menos tan grande como lo era en *X*. La mínima elipse capaz de hacerlo así es *F*. El centro de *F*, situado en *Y*, no es factible. Por ello, el algoritmo exige que la próxima elipse, *G* (*en negro*), cuyo centro es *Z*, contenga todos los puntos de la elipse *F* situado del lado factible de la restricción $4A \times 4B = 160$ (*en color*), que no es respetada por el punto *Y*. El proceso continúa, recortando la elipse recién construida con un contorno de la función objetivo cuando el centro de la elipse es factible; si no lo fuera, la elipse sucesiva habría de circunscribir la parte de la elipse recién construida que cumpliera las restricciones violadas por su centro. Las áreas de las elipses se contraen con rapidez y garantizan que sus centros converjan hacia una solución óptima.

programa de producción le dará mejores resultados. En el ejemplo considerado hay cinco vértices. El situado en (12,28), que representa la producción de 12 barriles de *ale* y 28 de cerveza, reporta unas ganancias (12 × 13 dólares) + (28 × 23 dólares), es decir, 800 dólares. Y éste será el máximo beneficio que nuestro cervecero pueda obtener.



LA PROPIEDAD DE CONVEXIDAD de la programación lineal estipula que, para cualquiera dos puntos de la región factible, el segmento rectilíneo que los conecta tiene que estar totalmente contenido en la región. Unicamente la región factible de la izquierda es convexa. La región del diagrama de la derecha constituye un conjunto de puntos aislados. La convexidad de la región factible garantiza que todo máximo local de una función objetivo lineal sea también máximo global. (Ilustraciones realizadas por Jerome Kuhl.)

La inclusión de restricciones adicionales (por ejemplo, escasez de levadura) no modificaría en sustancia la interpretación geométrica del problema del cervecero. A lo más, el polígono podría ahora tener seis lados, y no cinco. Por otra parte, de ser posible elaborar un tercer producto, las consecuencias geométricas tendrían mayor importancia, pues, ahora, el modelo geométrico sería tridimensional. Con tres variables de decisión, las desigualdades definen semiespacios limitados por planos, en lugar de semiplanos limitados por rectas, como sucedía en el caso de sólo dos variables. La región factible ya no es un polígono; más bien parecería una gema tallada, un politopo tridimensional, cuyas caras serían todas poligonales. Al aumentar el número n de variables de decisión, la interpretación geométrica sigue en esencia siendo válida, pero cada vez resulta más difícil imaginar el politopo n-dimensional correspondiente a la región factible, cuyas "caras" estarían contenidas en hiperplanos $(n - 1)$-dimensionales. Empero, los vértices continúan exhibiendo las mismas propiedades que antes, y su situación en el hiperespacio puede determinarse por medios de álgebra lineal, que relevan de su función a la intuición geométrica.

Pudiera parecer que siendo necesario calcular tan sólo los valores de la función objetivo correspondientes a los vértices de la región factible, sería posible resolver prácticamente el problema de programación lineal por simple enumeración. Sin embargo, como ya ocurría en el problema de asignación, el número de posibilidades a examinar crece vertiginosamente. Un problema con 35 variables de decisión y 35 restricciones nunca podría resolverse en la práctica por semejante método.

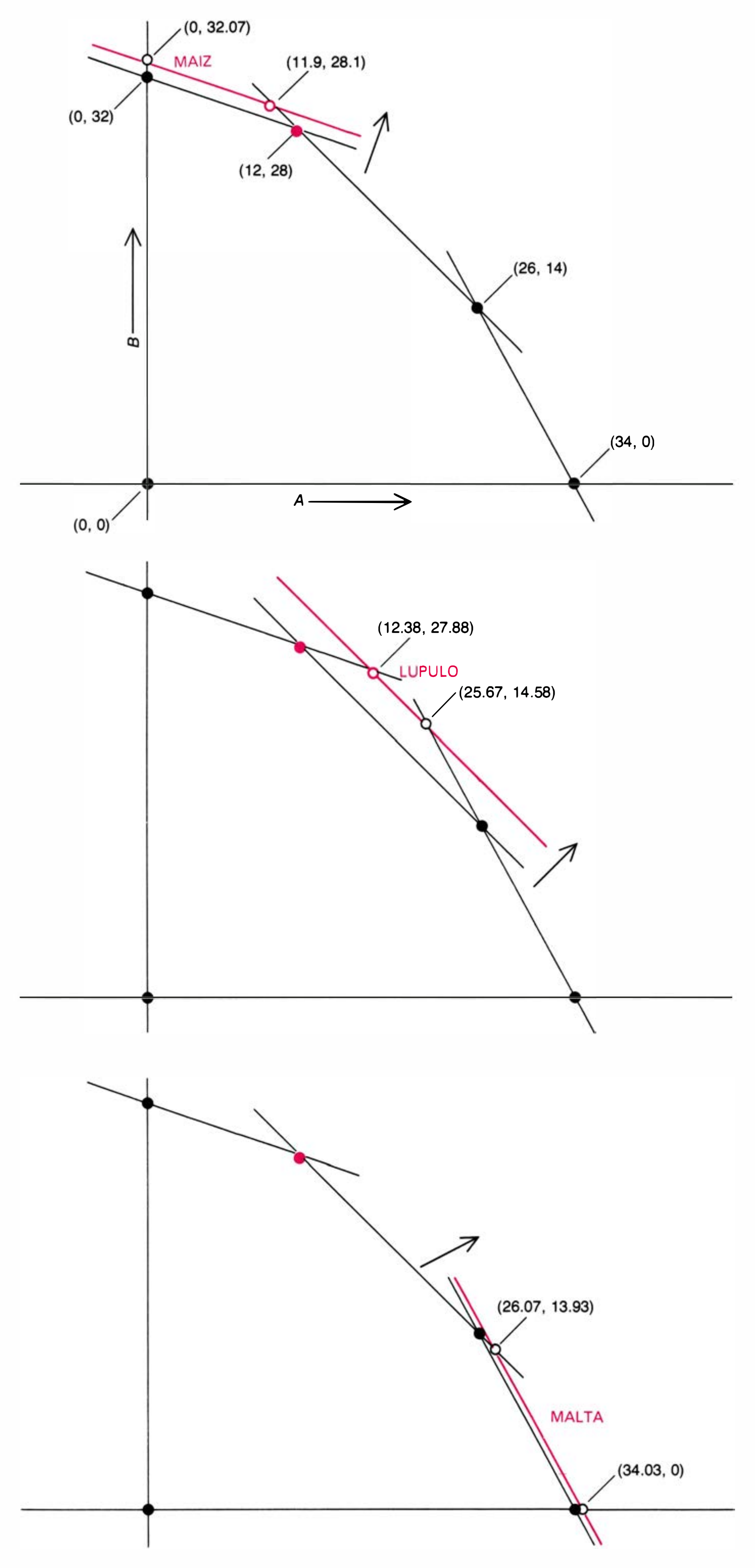
En el método del símplex, ideado por Dantzig, se estudian los valores de la función objetivo en los vértices de la región factible, pero no de manera exhaustiva, sino por un proceso selectivo. En el caso de nuestra fábrica de cervezas, el método podría empezar por el vértice situado en el origen (0,0). En él no hay producción, y las ganancias son nulas. Prosiguiendo a lo largo de uno de los lados que parten del origen se alcanzan puntos donde la función objetivo toma valores positivos. En el método del símplex se selecciona uno de tales lados, el eje B, pongamos por caso, que es recorrido hasta su otro extremo, el vértice (0,32). El plan de producción correspondiente consistiría en fabricar 32 barriles de cerveza y nada de *ale*, logrando así un beneficio de 736 dólares. De este vértice parte otro lado, conducente a valores aún más grandes de la función objetivo. El algoritmo del símplex nos lleva así al vértice (12,28), situado en el extremo del segundo tramo. En este caso se producirían 12 barriles de *ale* y 28 de cerveza, con beneficio de 800 dólares. Todos los lados concurrentes en (12,28) llevan a vértices menos ventajosos. Por consiguiente, el algoritmo se detiene en él, declarando que (12,28) es una solución óptima (que pudiera no ser única).

En general, el método del símplex va recorriendo las aristas de un politopo, explorando tras cada vértice otro adyacente, buscando siempre mejorar el valor de la función objetivo. El proceso puede arrancar de un vértice cualquiera, y se detiene cuando ninguno de los vértices contiguos al último explorado consigue mejorar el valor de la función objetivo recién calculado. La validez de esta regla de detención se debe exclusivamente a que la región factible es convexa: es la convexidad lo que garantiza que los vértices donde una función lineal alcanza máximos locales sean también máximos globales. Por ello, para saber si es o no posible mejorar los valores ya alcanzados es suficiente estudiar la vecindad de cada vértice.

¿Cómo podemos saber si uno de los lados que parten de un vértice nos permitirá mejorar el valor de la función objetivo? Tenemos la clave en la noción de "valor marginal", que es un "valor" asignado en cada vértice a cada uno de los recursos escasos. Para mejor comprender el método algebraico ordinariamente empleado para ir de un vértice a otro en los problemas de programación lineal con muchas variables, conviene examinar primero el significado e importancia de los valores marginales.

Volvamos al problema de la cervecería. En el vértice correspondiente al máximo, el empresario se encuentra con 210 kilos de malta sobrante, no requeridos por el programa de producción correspondiente al vértice óptimo. Añadir o detraer un kilo de malta de las existencias iniciales –1190 kilos– no modificaría la ganancia esperada. Por otra parte, de disponer de otra unidad más de lúpulo, sería posible incrementar en 2 dólares la ganancia total. Este incremento en las ganancias por unidad de ingrediente es el valor marginal del lúpulo en el vértice óptimo; podemos interpretar que tal valor traduce el efecto de "alejar" del origen la recta que expresa la escasez de lúpulo, de forma que ahora reflejase la disponibilidad de una unidad extra de este ingrediente [*véase la ilustración de la página siguiente*]. Análogamente, el valor marginal del maíz es, en este vértice, de 1 dólar por kilo.

Los precios marginales admiten interpretación económica natural. Si el cervecero pudiera lograr una unidad más de lúpulo, sus ganancias aumentarían en 2 dólares. Por tanto, de haber oferta de lúpulo a menos de 2 dólares por unidad, valdría la pena adquirir algunas unidades más. Por otra parte, si hubiera un comprador que ofertase más de 2 dólares por el lúpulo, al cervecero podría convenirle sacrificar en



parte su producción de cerveza y *ale* y vender algo de su lúpulo. No queremos decir que la compra o venta de este ingrediente deba realizarse indefinidamente a 2 dólares por unidad; sino que mientras las compras no excedan de 19 unidades, ni las ventas de 32, el precio de 2 dólares por unidad será quien decida en este ejemplo la conveniencia de realizar una u otra cosa.

Los precios marginales se llaman también precios imputados, pues indican el peso relativo con que cada uno de los ingredientes o recursos escasos contribuye a la rentabilidad de cada uno de los productos elaborados. Por ejemplo, cada barril de *ale* precisa cinco kilos de maíz, a precio imputado de 1 dólar por kilo; cuatro unidades de lúpulo, al precio imputado de 2 dólares cada una, y 35 kilos de malta, cuyo precio marginal es nulo. El precio imputado total del barril de *ale* es igual a 13 dólares, el beneficio esperado por él.

Supongamos que pueda fabricarse también cerveza de poca graduación, del tipo llamado *lager.* Cada barril de esta cerveza ligera consume dos kilos de maíz, cinco medidas de lúpulo y 24 kilos de malta. ¿Qué beneficios deberían obtenerse en cada barril de este nuevo producto para que estuviera justificado distraer recursos de la elaboración de cerveza y *ale*? Los precios imputados a los ingredientes resuelven esta cuestión. El precio total imputado de los ingredientes de la cerveza ligera es, en total, (2 × 1 dólar) + (5 × 2 dólares) + (24 × 0 dólares), es decir, 12 dólares. Esta cifra expresa las ganancias que dejarían de obtenerse al distraer recursos necesarios para fabricar cerveza y *ale* y consumirlos en cerveza ligera. Por consiguiente, para que valga la pena preparar este nuevo brebaje, será necesario que cada barril reporte al menos 12 dólares de beneficio.

El papel de los precios marginales a

INCREMENTOS MARGINALES de las existencias de recursos escasos; modifican de forma previsible los beneficios potenciales de nuestro hipotético fabricante de cerveza. De disponer de un kilo más de maíz, el beneficio máximo aumentaría en un dólar. El correspondiente cambio de valor de la función objetivo se refleja en el desplazamiento del vértice óptimo de la región factible, desde el punto de color al círculo de igual tono (*gráfico superior*). Si dispusiéramos de más lúpulo, el beneficio máximo aumentaría en dos dólares (*gráfico central*). Pequeñas modificaciones en las existencias de malta no modificarían el beneficio alcanzable, pues la malta es ya recurso excedentario (*gráfico inferior*). Los incrementos de beneficios implicados por incrementos de una unidad en los recursos disponibles se denominan precios imputados o precios fantasmas, y su función es dirigir el algoritmo de un vértice a otro. Para una mayor claridad, los incrementos marginales de las figuras están exagerados.

la hora de enjuiciar la conveniencia de elaborar nuevos productos puede ayudarnos a explicar la forma en que el método del símplex va determinando qué aristas del politopo nos llevarán desde un vértice a otro adyacente con mejores valores de la función objetivo. Supongamos que en este momento estamos examinando el vértice (0,32) en el problema de la cervecería. Podemos primero averiguar el precio marginal de cada recurso en este programa de producción. Hay en este caso malta y lúpulo en exceso; sus valores marginales son por tanto nulos. Sin embargo, el maíz es escaso, y así pues, su valor marginal es positivo. Puesto que en (0,32) estamos preparando solamente cerveza, el valor marginal de los ingredientes precisos por barril de cerveza debe ser, en conjunto, igual al benecificio que con él se consiga, esto es, 23 dólares. Como sólo el maíz tiene aquí valor marginal no nulo, y puesto que en cada barril de cerveza se consumen 15 kilos, el valor marginal del maíz en este plan de producción es de 23 dólares divididos entre 15 kilos, o sea, aproximadamente 1,53 dólares por kilo.

¿A cuánto asciende el valor marginal de los ingredientes de un barril de *ale* en el vértice (0,32)? Si tal montante fuese inferior al beneficio alcanzable con la venta de un barril de esta bebida, el cervecero podría aumentar el total de ganancias dedicando algunos de los recursos necesarios para producción de cerveza a la elaboración de *ale*. Fabricar un barril de *ale* sólo requiere cinco kilos de maíz. Por consiguiente, el valor marginal de los ingredientes del barril de *ale* es aquí de (23 dólares / 15) × 5, aproximadamente, 7,67 dólares. Como sucedía en el ejemplo de la cerveza ligera, el valor marginal representaría la disminución de beneficios resultante de dedicar recursos necesarios para fabricar cerveza a la producción de *ale*. En vista de que tal pérdida es menor que los 13 dólares de ganancia esperada en la venta de cada barril de *ale*, será ventajoso preparar más *ale* de lo especificado en el programa correspondiente al punto (0,32). En efecto, manteniendo constante el consumo total de maíz, y dedicando parte del maíz a la producción de *ale* a costa de la cerveza, el cervecero puede aumentar sus beneficios a razón de 13 − 7,67 = 5,33 dólares por cada barril de *ale*. Puesto que la cantidad de maíz permanece constante, producir más *ale* equivale a desplazarse a lo largo de un lado de la región factible, a saber, el que representa la limitación sobre existencias de maíz. Es precisamente de esta forma como el método del símplex detecta la posibilidad de mejorar la función objetivo recorriendo este lado. En este caso, el lado conecta el vértice (0,32) con el vértice (12,28) del polígono.

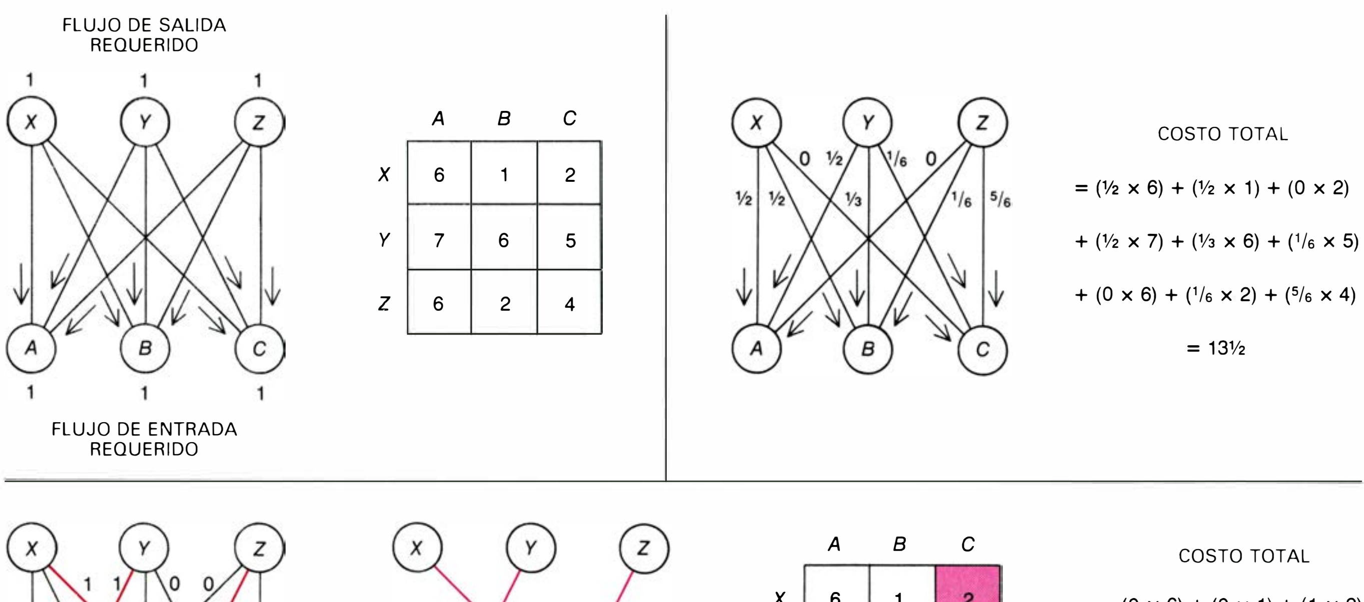
Incluso en problemas de programación lineal pequeños, la región factible puede tener enorme número de vértices, y cabe la posibilidad de que el método del símplex requiera, para resolver el problema, un número astronómico de iteraciones, esto es, de saltos de un vértice a otro adyacente. Por lo común, sin embargo, el método, en su búsqueda selectiva solamente explora una fracción casi despreciable del total de vértices. Quienes trabajan de forma habitual en la resolución de problemas grandes de programación lineal, con 2000 o 3000 restricciones y 10.000 o 15.000 variables de decisión, han observado que tales problemas suelen ocupar entre algunos minutos y varias horas a los ordenadores de gran velocidad. Sin duda, sería deseable poder enunciar con más precisión cuál es la eficiencia del método de símplex.

Una forma de convencer a los escépticos sobre la eficiencia de un algoritmo podría ser ofrecerles alguna clase de garantía. Tal garantía podría consistir en una función G más una declaración, que estipulase que cuando el problema tiene tamaño t el tiempo necesario para procesar el algoritmo hasta su término no excederá del valor $G(t)$. El tamaño t de un problema es una medida de la cantidad de información (frecuentemente expresada en número de *bits,* esto es, de dígitos binarios) requerida para especificar todos los datos de entrada. Este análisis corresponde al caso "más desfavorable", pues el valor de garantía está determinado por el más penoso que entre los problemas de tal tipo y tamaño t tenga que resolver el algoritmo. El tiempo de funcionamiento materialmente necesario para resolver el problema dependerá del ordenador concreto que vaya a utilizarse. Con objeto de independizar el valor de $G(t)$ del tipo de máquina utilizada, es costumbre expresarlo no en horas ni minutos, sino como número máximo de operaciones elementales (tales como sumas, multiplicaciones y comparaciones de valores) que tendría que efectuar el algoritmo.

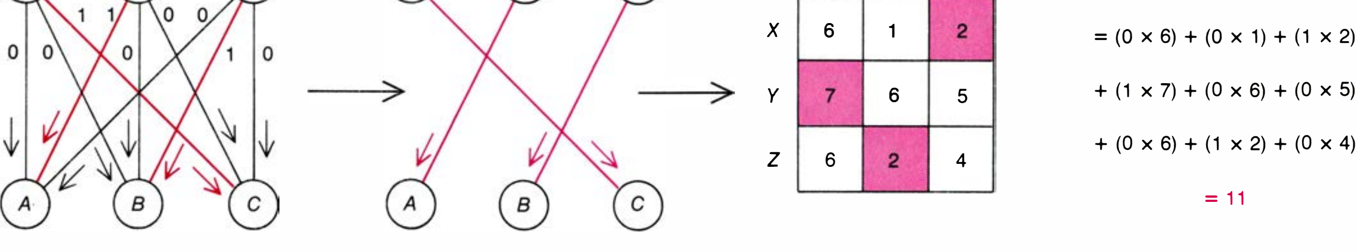
Cuando el valor de $G(t)$ es adecuadamente pequeño en relación con los valores de t que será necesario manejar, nos sentimos inclinados a reconocer que el algoritmo propuesto es eficiente. Por ejemplo, Harold W. Kuhn, de la Universidad de Princeton, ha ideado un algoritmo capaz de resolver el problema de asignación de tamaño n por n con no más de cn^3 operaciones elementales, siendo c una constante pequeña. Puede suceder que cuando n tome los valores 2, 3, 4 o 5, tal algoritmo sea más lento que la pura enumeración (si bien ningún ordenador necesitará más de un instante para resolver por cualquier método un problema tan pequeño). Mas cuando n se sale de tan estrecho campo, la $n!$, que mide el tamaño del proceso de enumeración, muy pronto deja atrás a cn^3. En realidad, la $n!$ acelera tan rápidamente su crecimiento –"explosivamente", podríamos decir– que, como ya he comentado, ninguno de los ordenadores existentes podría ejecutar los cálculos necesarios. El crecimiento de n^3 es mucho más moderado. Si n vale 35, el algoritmo de Kuhn aún puede resolver el problema en el tiempo de un guiño. Además, cuando n pasa de 35 a 36, el trabajo de cálculo correspondiente al algoritmo de Kuhn aumenta en el factor $(36/35)^3$, aproximadamente, 1,09. En el algoritmo de enumeración, el trabajo de cálculo queda, en cambio, multiplicado por 36.

En teoría de computación hay interés especial por los algoritmos llamados "de acotación polinómica", a saber, algoritmos donde G viene dada por una función polinómica de t. Una función polinómica es suma de varios términos; en cada uno de ellos figura t elevada a cierta potencia y multiplicada por un coeficiente constante. Para valores de t suficientemente grandes, todo algoritmo cuyo caso menos favorable requiera un número de operaciones del orden de $t!$, o de una función exponencial, como 2^t, será con certeza más lento que cualquier algoritmo polinómicamente acotado. Por ello, en teoría de computabilidad la acotación polinómica suele considerarse como criterio teórico de expresión de eficiencia de los algoritmos.

En vista de que el método del símplex se desenvuelve bien al resolver problemas prácticos, podríamos conjeturar que se trata de un algoritmo eficiente (polinómicamente acotado). Más todavía: los profesionales que lo utilizan han descubierto, por vía empírica, que el número de iteraciones que el algoritmo debe ejecutar es sensiblemente función lineal del número de restricciones. Por sorprendente que nos parezca, el método del símplex no está polinómicamente acotado. Victor La Rue Klee, Jr., de la Universidad de Washington, y George J. Minty, Jr., de



	A	B	C
X	6	1	2
Y	7	6	5
Z	6	2	4



	A	B	C
X	6	1	2
Y	7	6	5
Z	6	2	4

PROBLEMAS DE FLUJOS en redes; constituyen casos particulares de un tipo especial de problemas, llamados de "programación entera", fácilmente resueltos por técnicas de programación lineal. En los problemas de programación entera, las variables de decisión sólo pueden tomar valores enteros. En general, tales problemas pertenecen al conjunto *NP*-completo, y son considerados difíciles. El costo de cada enlace de la red (*diagrama superior izquierdo*) viene dado en la matriz numérica asociada. Observemos que tal matriz es idéntica a la posible matriz de un problema de asignación de tamaño 3 por 3. Sin embargo, las soluciones del problema de asignación han de tomar valores enteros, y las restricciones sobre los flujos no bastan para garantizarlo así. Hay muchos flujos de valores fraccionarios capaces de cumplir todas las restricciones (*diagrama superior derecho*). Empero, las soluciones extremales de los problemas de flujo determinadas por medio de programación lineal son siempre enteras, a condición de que los flujos que entran o salen de cada nodo sean todos de valor entero. Por tanto, el problema de asignación puede ser interpretado como caso particular del problema de flujos de una red (*abajo*).

la de Indiana, han construido una familia infinita de problemas de programación lineal en los cuales el método del símplex es sustancialmente tan poco eficaz como la pura enumeración.

No deja de ser paradójico que un algoritmo tan eficiente en la práctica pueda no ser considerado así desde un enfoque teórico. Tengamos presente, sin embargo, que el criterio teórico analiza el rendimiento en los casos más desfavorables. Al parecer, los ejemplos que Klee y Minty han sido primeros en construir no se presentan en la práctica, o no lo han hecho hasta ahora. Con objeto de explicar el divorcio entre la gran eficiencia del método del símplex en la práctica y su pésimo rendimiento en los casos más desfavorables, los investigadores han intentado demostrar que, en cierto sentido probabilístico, los casos de bajo rendimiento son sumamente excepcionales, "patológicos" en su jerga. Las dificultades de semejante proyecto son considerables, si bien Dantzig ha dado cuenta recientemente de algunos interesantes progresos en esta dirección.

En el método del símplex se dispone de cierta flexibilidad al ejecutar cada paso de iteración, es decir, hay cierta libertad de elección para ir de un vértice a otro de los adyacentes en el caso de que todos ellos mejoren el valor del objetivo. Las construcciones de Klee y Minty ponen de manifiesto que una de las estrategias de elección más difundidas genera algoritmos no acotados polinómicamente. Otros investigadores han construido después contraejemplos parecidos para otras estrategias. Se desconoce, empero, si habrá alguna traducción algorítmica del método del símplex que admita acotación polinómica. Más todavía: hasta muy recientemente se ignoraba si los problemas de programación lineal admitirían algún tipo de algoritmo acotado polinómicamente; mientras permaneció abierto, este problema tuvo gran interés teórico. Finalmente, L. G. Khachian, de la Academia de Ciencias de la Unión Soviética, demostró que un algoritmo ideado por tres matemáticos rusos, N. Z. Shor, D. B. Iudin y A. S. Nemirovskii, podía ser ejecutado en tiempos de acotación polinómica.

Debido a que el algoritmo de Shor, Iudin y Nemirovskii no presupone que las restricciones o el objetivo hayan de ser lineales, el nuevo algoritmo difiere fundamentalmente del método del símplex. Llamado "método del elipsoide", en él se comienza tomando un gran elipsoide centrado en el origen. Por lo común este elipsoide inicial es una esfera de radio lo bastante grande como para garantizar que en ella está contenida la solución óptima. El algoritmo procede entonces a construir elipsoides cada vez más pequeños, en los que también es seguro estará contenida la solución óptima. El volumen de los sucesivos elipsoides se contrae lo bastante rápidamente como para garantizar que sus centros vayan convergiendo hacia la solución óptima. Para emplear el algoritmo no es preciso dibujar los elipsoides; su descripción se logra a través de sencillas fórmulas iterativas. Al analizar el número de operaciones necesarias para construir un nuevo elipsoide a partir de su predecesor, tal número admite una cota del orden de $(m + n)n^6l$, donde m denota el número de restricciones, n es el de variables de decisión y l es el logaritmo de base 2 del máximo coeficiente que figure entre los datos del problema.

Desde luego, tal expresión es una acotación polinómica, pero lo mismo que cualquier otra acotación de este ti-

po, puede tomar valores enormes si *m*, *n* y *l* alcanzan valores suficientemente grandes. ¿Pudiera suceder, como es el caso del símplex, que su comportamiento práctico fuese muy superior al predicho por la acotación para el caso menos favorable? Los primeros resultados son más bien descorazonadores; los experimentos de cómputo efectivo realizados han dado muestras de que la convergencia hacia la solución es, con el método del elipsoide, extremadamente lenta. Hay familias de problemas de sencillez notoria capaces de forzar al algoritmo a tiempos de funcionamiento iguales a los máximos predichos por la cota, en contraste con la complicada expresión de los problemas ideados para vencer al método del símplex.

Otra de las principales dificultades del método del elipsoide concierne a la cantidad de memoria necesaria para procesarlo. En la práctica, los problemas grandes de programación lineal tienden a ser de "baja concentración": la proporción de coeficientes no nulos que figuran en ellos suele ser pequeña. Al llevar a la práctica el método del símplex se puede conseguir que la baja concentración se mantenga durante el proceso de datos; y esta propiedad es de la máxima importancia, pues en la memoria del ordenador es preciso cargar tan sólo los datos no nulos. No hay a la vista ningún procedimiento para sacar ventaja de esta propiedad cuando se utiliza el método del elipsoide [*véase la ilustración de la página 70*].

Aunque las expectativas de que el método del elipsoide llegue a cobrar importancia práctica son más bien pesimistas, es preciso aguardar el desarrollo de los acontecimientos. En el breve lapso transcurrido desde la publicación de la nota de Khachian se han realizado esfuerzos considerables que han aportado mejoras de importancia. La dificultad de la tarea parece, sin embargo, formidable.

La programación lineal es un instrumento de gran importancia práctica y no menor interés teórico. Para el matemático, se trata de una teoría de inecuaciones lineales, elegante contrapunto de la clásica teoría de ecuaciones de primer grado. Para los encargados de tareas de planificación, tanto gubernamental como empresarial, sirve de preciosa ayuda para tomar decisiones o formular proyectos de gran alcance. Tal vez la más importante consecuencia del logro teórico de Khachian sea aumentar el interés por los aspectos, tanto prácticos como teóricos, de la programación lineal.



ENJAMBRE DE ABEJAS MELIFERAS que se forma en la rama de un árbol. Un enjambre típico consta de una reina y varios miles de abejas, que vienen a representar la mitad de la población de la colmena que han abandonado súbitamente a finales de la primavera para establecer una nueva colonia. La forma subcónica y el estrecho apretujamiento de las abejas de este enjambre son característicos y los mantendrán hasta dar con un nuevo emplazamiento.

Regulación de la temperatura en el enjambre de abejas melíferas

Para volar, las abejas melíferas deben entrar en calor. De ahí que la temperatura de un enjambre inmóvil deba estar estrechamente controlada, para que, cuando las exploradoras den con un nuevo lugar, el enjambre esté listo para despegar

Bernd Heinrich

El enjambramiento de las abejas melíferas es un fenómeno que ocurre a finales de la primavera, cuando una reina y varios miles de abejas abandonan súbitamente la colmena, dejando en ella a otra reina y al resto de las abejas residentes. En cosa de horas, o días a lo sumo, el enjambre se instala en un nuevo lugar. Pero en el ínterin las abejas se comportan de un modo extraño. El enjambre se aglutina, constituyendo una masa en forma de barba, sobre la rama de un árbol o algún otro objeto firme. Permanece allí prácticamente inmóvil, salvo unas cuantas exploradoras que salen en busca de un lugar adecuado para vivir, como un árbol hueco. Si se sacude el enjambre, la mayoría de abejas caerán al suelo. Están demasiado frías para volar: sus músculos de vuelo no pueden funcionar hasta que hayan entrado en calor. Pero cuando llega el momento de desplazarse al nuevo hogar, el enjambre apenas tarda un minuto en remontar el vuelo. Aparentemente, el enjambre regula de algún modo su temperatura. He investigado este tema en enjambres libres y cautivos y he llegado a la conclusión de que las abejas poseen un sistema de termorregulación notablemente ajustado. Si se piensa en ello, es evidente que en un clima de zona templada un sistema así es esencial para que el enjambre consiga su fin último: establecer una colmena viable.

La función primaria de la enjambrazón en la abeja melífera "europea", *Apis mellifera* (que se exportó a Norteamérica desde Europa), es la reproducción de las colonias. Por lo general, la reina y aproximadamente la mitad de las obreras de una colonia abandonan la colmena cuando ésta deviene superpoblada. Detrás queda una reina hija.

Un fenómeno que guarda cierta relación es la "fuga": toda la colonia abandona su domicilio si éste se torna inhóspito. Algunas especies se fugan más fácilmente que otras. La diferencia puede residir en la adaptación de las abejas a su clima. En los trópicos asiáticos, la melífera gigante (*Apis dorsata*) y la melífera enana (*A. florea*), que normalmente construyen sus panales colgantes al aire libre, se fugan con frecuencia, probablemente con el fin de desplazarse a zonas donde el alimento sea más abundante. La tercera melífera tropical, la abeja india (*A. indica* o *A. cerana*), construye sus panales en cavidades y también se fuga fácilmente, al igual que la raza africana de *A. mellifera*, la "abeja asesina", que, desde su introducción en Sudamérica, se ha extendido hacia el norte a una velocidad de unos 50 kilómetros anuales.

El clima cálido explica por qué las tres especies tropicales y las razas meridionales de *A. mellifera* se fugan con tanta facilidad. Las colonias tienen mucho que ganar si sacan partido de recursos alimentarios variables, y no precisan acumular grandes reservas de alimento para sobrevivir a los inviernos de la zona templada. Movilidad equivale a viajar deprisa y a no perder mucho tiempo escogiendo un lugar para el nido.

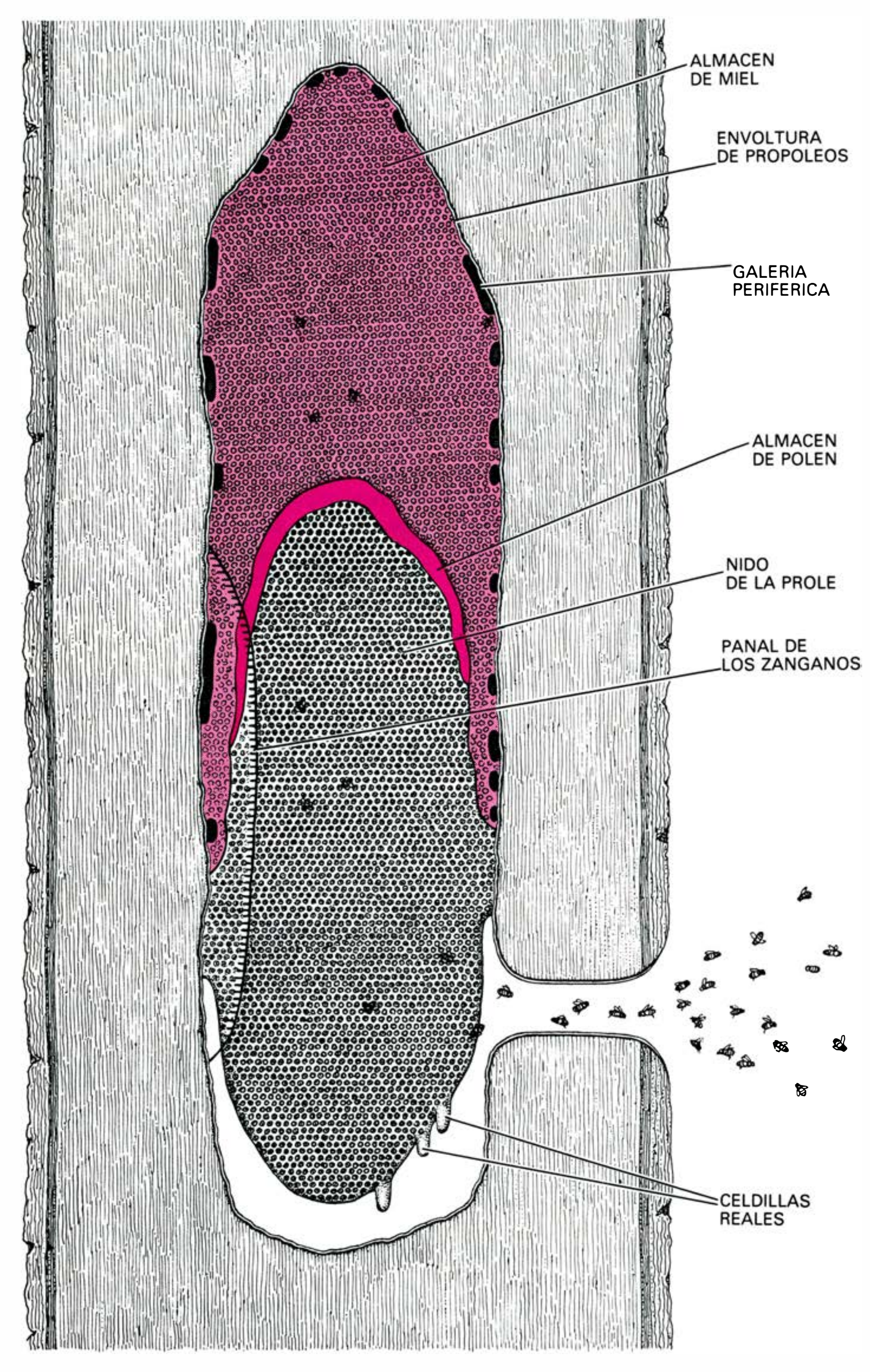
Sin embargo, en la zona templada septentrional, la abeja melífera tiene que elegir muy bien el tipo de cavidad que ocupa. La abertura ha de ser pequeña, y el tamaño de la cavidad debe ser tal que el enjambre pueda controlar su temperatura durante el invierno. Además, después de ocupar una cavidad, las abejas acumulan cantidades de alimento mucho mayores que las que pueden llevar consigo, pues el alimento es, de hecho, el combustible que caldea la colmena. Por tanto, las abejas deben disponer de las respuestas sociales necesarias para administrar con la máxima eficiencia sus recursos alimentarios. No podrán abandonar el nido o la zona cuando las condiciones son peores, en invierno. Estas abejas están adaptadas a quedarse, no a marcharse. Su enjambrazón está más dirigida a la reproducción que a buscar un mejor alojamiento.

No hay cantidad de alimento almacenado ni respuesta social que acondicione el ambiente interno de la colmena que resistan, por mucho tiempo, el viento y el frío del invierno septentrional si las abejas no se hallan en un refugio adecuado. De ahí que la búsqueda, la evaluación de domicilios adecuados y la competencia por ocuparlos sean factores básicos en la biología de la abeja melífera de la zona templada septentrional. Ello requiere distintas adaptaciones de comportamiento; la termorregulación es, en todas ellas, de capital importancia. Vale la pena considerar las adaptaciones como telón de fondo de la termorregulación.

Un importante prerrequisito para tener éxito en la enjambrazón es la correcta medición del tiempo. La razón por la que los enjambres suelen salir de las colmenas en primavera se debe a que, entonces, las abejas disponen de más tiempo para encontrar un lugar apropiado donde vivir, ocuparlo, llenarlo de bresca y abastecer el panal con acopio de miel para el invierno siguiente.

Mientras la reina y la mayoría de las demás abejas de un enjambre recién formado cuelgan casi inmóviles de algún objeto, por lo general próximo a la

colmena que acaban de dejar, varias abejas exploradoras comienzan la búsqueda de una nueva morada. Constituyen la reducida parte (menos del 5 por ciento) de la población total que parece ocuparse de la exploración y no del trabajo de la colmena. En una colmena asentada, las exploradoras suelen ser las abejas que buscan nuevas fuentes de alimento, como plantas en floración y colmenas débiles donde robar. En el enjambre, son éstas las abejas que buscan posibles emplazamientos para la colmena.



EMPLAZAMIENTO TIPICO DE UN NIDO, en un árbol parcialmente hueco, ilustrado sobre la base de varios nidos naturales inspeccionados por Thomas D. Seeley, de la Universidad de Yale, y Roger A. Morse, de la Universidad Cornell. Un lugar como éste es el que las abejas exploradoras examinan atenta y repetidamente mientras el enjambre aguarda en su característica forma de agrupamiento. Una vez las exploradoras han llegado a un acuerdo acerca del emplazamiento, alertan al enjambre, que puede emprender rápidamente el vuelo porque las abejas han mantenido la temperatura de su núcleo cerca de los 35 grados Celsius, el nivel óptimo para el vuelo. El enjambre, guiado por las exploradoras, vuela hasta el lugar, se instala rápidamente y comienza a construir las distintas estructuras que aquí se indican.

La importancia del tipo de emplazamiento que eligen se pondera en los estudios que Thomas D. Seeley, de la Universidad de Yale, ha hecho en enjambres cercanos a Ithaca (Nueva York). Este autor ha encontrado que la tasa de mortalidad más elevada en la colmena correspondía al invierno, tasa que estaba inversamente correlacionada con la calidad de la cavidad utilizada de nido. También ha descubierto que las abejas exploradoras prestan gran atención al tamaño de la abertura de un emplazamiento potencial para el nido, a la temperatura interna y a las corrientes de convección internas del mismo. Las exploradoras miden también las dimensiones interiores de la cavidad. Según las mediciones de Seeley, el volumen de la cavidad elegida como domicilio tiene por lo general entre 20 y 100 litros.

Martin Lindauer, de la Universidad de Würzburg, estudió enjambres de abejas en Munich, después de la segunda guerra mundial, cuando trabajaba en el Instituto Zoológico de la capital bávara. Encontró que la mayoría de los enjambres escogían sus cavidades en las ruinas, que por aquel entonces abundaban en la ciudad. Llegó hasta predecir qué colmena potencial ocuparía un determinado enjambre.

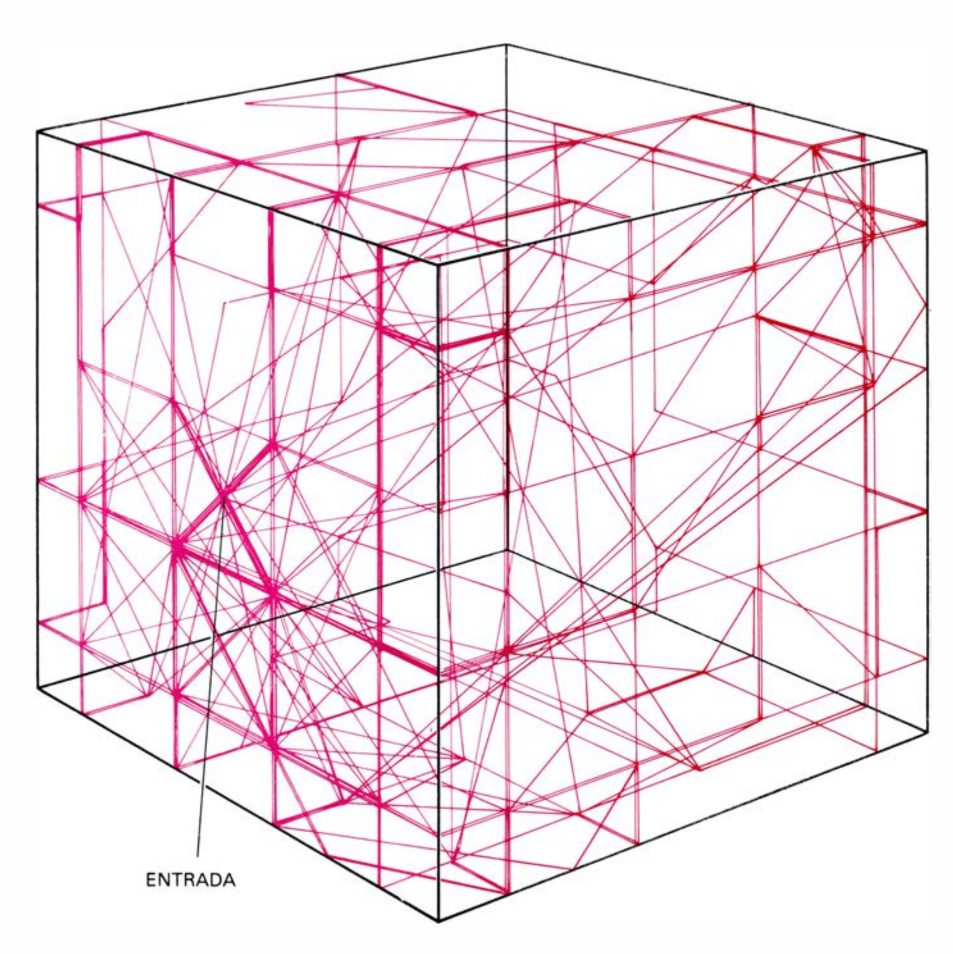
Lindauer observó que las abejas exploradoras, algunas de las cuales había marcado, empezaban, antes o después, a "danzar" sobre la superficie de un enjambre agregado, del mismo modo que las abejas que danzan en una colmena establecida indican la distancia y la dirección hacia fuentes potenciales de alimento. En ocasiones, varias exploradoras advertían simultáneamente acerca de distintas cavidades. Antes de iniciar la danza, una exploradora visitaba varias veces una cavidad; dejaba de señalar el emplazamiento si la cavidad se hacía demasiado húmeda o demasiado cálida, o bien cambiaba desfavorablemente en algún otro sentido. Las exploradoras también prestaban atención a las danzas de las demás, visitando los otros posibles emplazamientos para la morada que indicaban las danzas. Al final, todas las exploradoras se acercaban repetidamente al mismo lugar y advertían de él al enjambre.

Llegados a este punto, cabe preguntarse ¿cómo se despierta al enjambre y se le guía hasta el emplazamiento? Seeley y sus colegas observaron exploradoras marcadas que habían estado investigando una colmena vacía y danzaban sobre el enjambre. Media hora antes de

que el enjambre emprendiera el vuelo, las exploradoras efectuaron rápidos viajes zumbantes a la colmena. Estas carreras cortas parecen ser una señal de alerta, y no siempre van seguidas por el despegue del enjambre, lo que sugiere que la acción de remontar el vuelo de forma casi instantánea puede obedecer a otra señal.

Cuando un enjambre ha emprendido el vuelo, se desplaza lentamente al principio, en una formación cilíndrica de unos 10 metros de diámetro. La lentitud del vuelo quizá se deba a que las abejas están verificando el olor de la sustancia real, el ácido 9-oxodec-2-enoico. (Un enjambre puede emprender el vuelo sin una reina, pero no llegará lejos.) El enjambre acelera pronto hasta una velocidad de unos 11 kilómetros por hora, pero algunas abejas, las "veloces", se lanzan mucho más rápidamente a través del grupo en la dirección de la nueva cavidad. Lindauer y Seeley piensan que las veloces pueden ser las abejas exploradoras que señalan el camino.

Las exploradoras también indican el camino químicamente, en particular cuando la meta está cercana. Llegan a la entrada de la cavidad en cabeza de la masa de abejas y liberan olor procedente de las glándulas de Nassanoff, situadas en el extremo del abdomen. Parece que el olor actúa como un marcador final para el enjambre. En pocos minutos, los millares de abejas penetran en la cavidad. A las pocas horas están limpiando de restos el lugar, construyen panales y salen en busca de néctar y polen. Se ha establecido una nueva colonia.



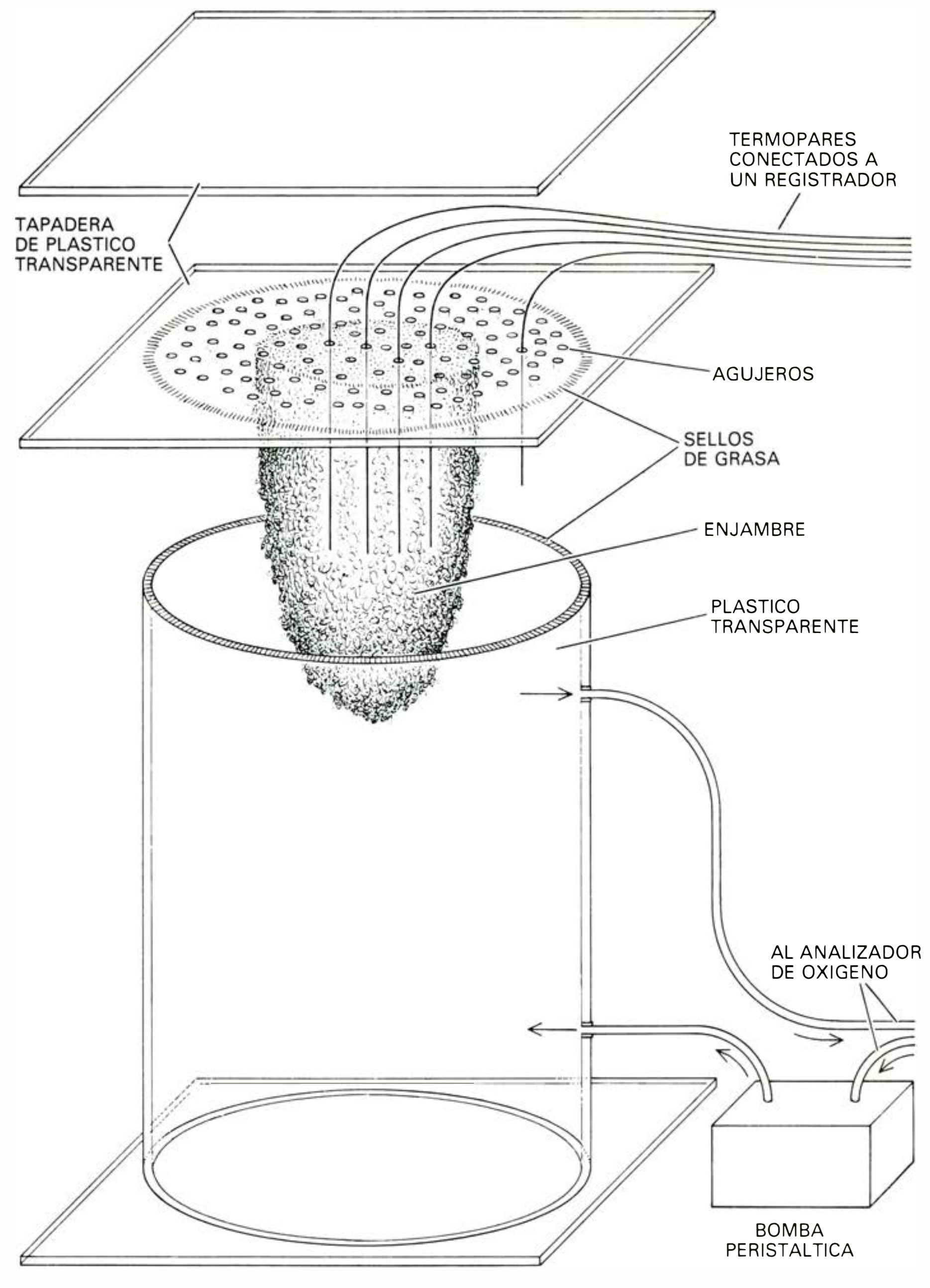
PAUTA DEL MOVIMIENTO de una única abeja exploradora en varias inspecciones de una caja cúbica. representada aquí de forma algo simplificada a partir de una reconstrucción realizada por Seeley. Las líneas de color indican las trayectorias de la abeja; para simplificar se han trazado rectas. Seeley empleó un sistema de rejillas que le permitió determinar la trayectoria de la posición de la abeja exploradora.

Si las exploradoras de un enjambre no consiguen encontrar una cavidad adecuada en unos días, o bien no llegan a un acuerdo acerca de cuál de las distintas alternativas es la mejor, las obreras comienzan a construir panales sobre el objeto del que penden. Cuanto mayor es el panal que construyen, menos probable es que abandonen el lugar, pues el panal representa una inversión sustancial de esfuerzo. Las colmenas abiertas pueden sobrevivir en lugares donde el invierno es suave, pero en la zona templada las abejas de un panal abierto mueren invariablemente durante el invierno, pues son incapaces de mantener el calor sin un refugio convenientemente aislado.

La medición de la temperatura en las colmenas data, por lo menos, del siglo pasado. Se sabe desde hace tiempo que la colmena funciona como una incubadora de larvas y pupas. Las abejas mantienen la temperatura de la prole a 35 o 36 grados Celsius. Cuando la temperatura del aire es baja, las abejas se amontonan sobre la prole y tiritan; la rápida contracción de sus músculos genera calor corporal. Cuando la temperatura del aire es alta, las abejas reducen la temperatura interna abanicando con las alas y mediante enfriamiento producido por evaporación. Si se deja que la prole se enfríe por debajo de los 30 grados C, las larvas o pupas mueren, o emergen con defectos de desarrollo, como alas deformadas. La temperatura no se controla rígidamente en aquellas secciones de la colmena que no contienen crías, por lo que se supone que la termorregulación la precisan, ante todo, las abejas inmaturas.

Un enjambre se compone enteramente de abejas adultas; no hay huevos, larvas, pupas o panales. En principio, por tanto, no tendría por qué haber termorregulación. Sin embargo, mis mediciones preliminares mostraban que el núcleo o centro del enjambre se mantenía generalmente casi a la misma temperatura a la que se halla la prole en una colmena, y que permanecía constante, o aumentaba, cuando disminuía la temperatura externa. Estas observaciones planteaban dos preguntas, que traté de responder. Era la primera: ¿cómo regulaban las abejas situadas en el centro del grupo su temperatura corporal con respecto a la temperatura externa, que varía, si, individualmente, sólo perciben la temperatura relativamente constante o en aumento del centro? Y la segunda: ¿cuál era el significado funcional de la termorregulación en el enjambre?

Capturé muchos enjambres tirándolos al suelo al sacudir las ramas en las que se encontraban. Me los llevé al laboratorio para determinar su temperatura y su tasa metabólica en grandes recipientes de respirometría. La tapadera de estos recipientes tiene pequeños agujeros, que sirven para la ventilación, de apoyo para las abejas (permitiéndoles formar un enjambre arracimado bajo la tapadera) y como aberturas para la inserción de instrumentos de medida. Para calcular la res-



APARATO EXPERIMENTAL diseñado por el autor para medir la temperatura y la tasa metabólica de enjambres cautivos. Aquí se muestra en una vista esquemática, con las partes separadas. Los agujeros de la tapadera inferior sirven de apoyo para las abejas y como fuente de ventilación. Mientras dura la medición, la tapadera continua se desliza sobre la tapadera de ventilación, cubriéndola, y se hace circular aire desde el sistema cerrado al analizador de oxígeno para determinar la tasa de consumo de energía de las abejas. Al mismo tiempo, el aparato permite medir la temperatura en distintas partes del enjambre.

piración cerré el recipiente con una segunda tapadera de plexiglás macizo colocada sobre la tapadera perforada; el aire del sistema sellado se hacía circular a través de un analizador de oxígeno para determinar la tasa de consumo de energía, es decir, la producción de calor.

Observé también enjambres vivos que se mantenían sobre el alféizar de una ventana exterior de mi laboratorio de la Universidad de California en Berkeley. En algunos de ellos había implantado una docena o más de termopares, cuyos cables entraban en el laboratorio; podía así registrar la lectura de cada termopar sin interrupción, durante horas o días, en un registro de papel continuo. Los enjambres libres respondían a las condiciones del medio, pero con los cautivos podía variar la temperatura reinante de cero a 30 grados C.

En todos los enjambres se apreciaba un fuerte gradiente térmico desde el centro hasta el manto, o perímetro. Las temperaturas más elevadas se encontraban cerca del centro. Tanto los enjambres pequeños, de 2000 abejas, como los que eran quince veces mayores mantenían una temperatura central de unos 35 grados C. Aunque la temperatura del núcleo variaba considerablemente, era independiente de la temperatura ambiente. Puesto que la tasa de enfriamiento pasivo en un enjambre es función de la superficie y de la temperatura ambiente, los resultados mostraban claramente que la temperatura del centro se regula de forma activa. Lo que no dejaba de sorprender, porque en el centro había sólo una fracción pequeña de la población total del enjambre.

El manto de los diferentes enjambres se mantenía por lo general a una temperatura mucho más baja. Si la temperatura ambiente descendía, la del manto también bajaba, permaneciendo de dos a tres grados por encima de aquélla. No obstante, esta pauta sólo aparecía cuando la temperatura ambiente era relativamente alta. Si descencía por debajo de los 20 grados C, la temperatura del manto se estabilizaba. Estos resultados mostraban que la temperatura del manto se regulaba de tal manera que permanecía cercana a la temperatura ambiente, pero no descendía por debajo de los 17 grados. La temperatura del resto del enjambre solía variar en un gradiente bastante suave entre el máximo de 36 grados del núcleo y el mínimo de 17 grados del manto.

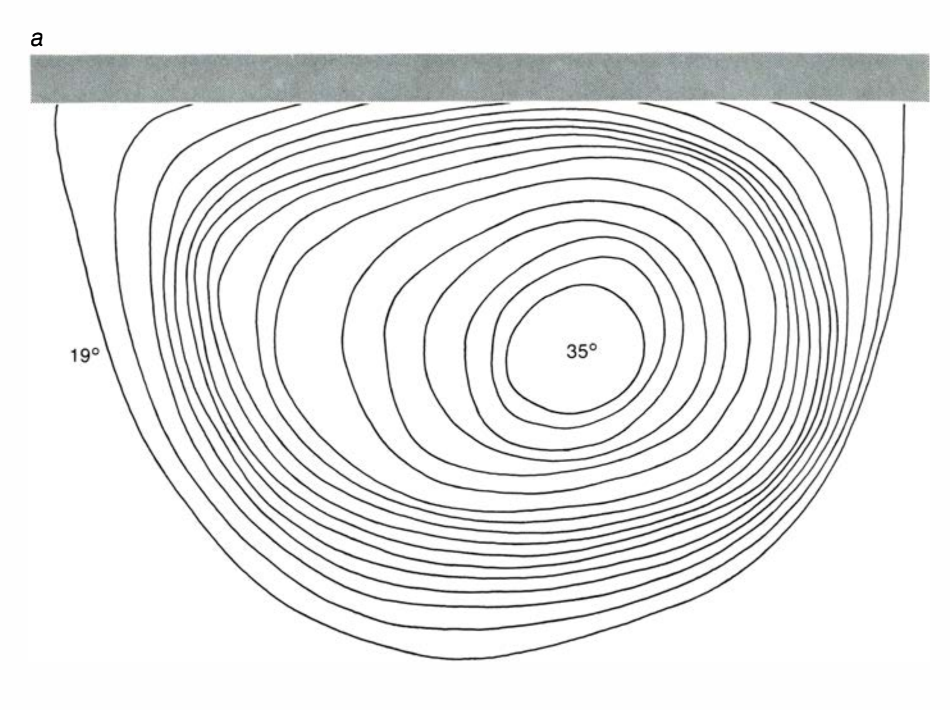
Este era el modelo general que resultaba de numerosas mediciones, pero también encontré variaciones y excepciones instructivas. Por ejemplo, los enjambres cautivos tenían con frecuencia temperaturas inferiores a las de los libres, presumiblemente porque en un enjambre cautivo existe menos actividad. Cuando se mantenía un enjambre pequeño a cinco grados C o menos durante varios días, las abejas colgaban en cadenas, aparentemente sin vida, del racimo fuertemente contraído. La temperatura del cuerpo de estas abejas se aproximaba a la del ambiente; las abejas terminaban por caerse. La mayoría de ellas presentaban el buche de la miel vacío, posiblemente porque habían agotado sus reservas de alimento al tiritar. Las abejas que persistían en el manto del enjambre, manteniendo una temperatura corporal de unos 17 grados C, podían arrastrarse, pero no volar.

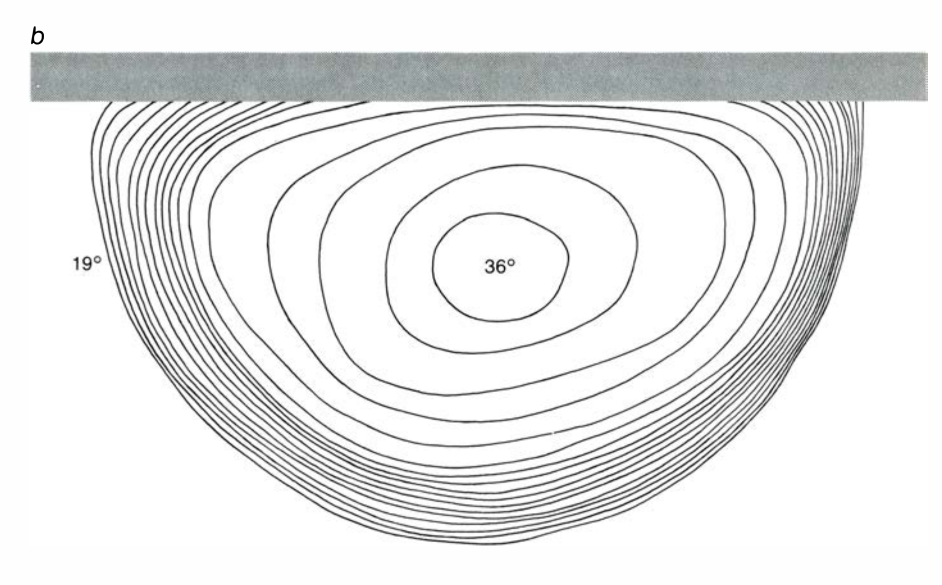
Las abejas de un enjambre libre pasaban el día en constante ida y venida, manteniéndose el manto a una temperatura notablemente más elevada que el de un enjambre cautivo. La tempera-

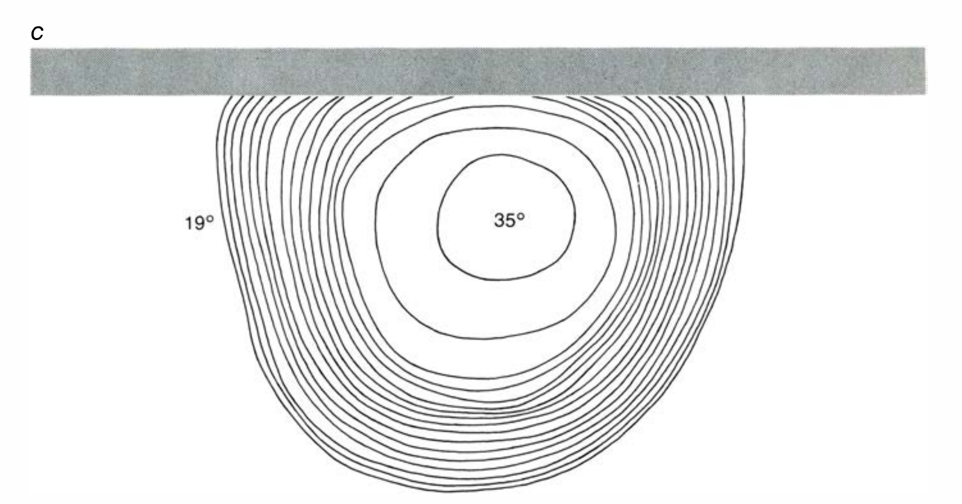
tura del manto de los enjambres libres fluctuaba durante el día por encima del mínimo. En un enjambre libre a punto de emprender el vuelo, sin embargo, el gradiente de temperatura desaparecía, y el manto alcanzaba la temperatura del núcleo, 36 grados C.

¿Cómo se regula la temperatura del centro? Parece razonable suponer que cuando la temperatura externa desciende, o cuando todo el enjambre está caldeándose para el vuelo, el enjambre, en su conjunto, consume más energía para contrarrestar la creciente pérdida de calor. Si ocurre así, un enjambre pequeño debe tener una tasa metabólica superior a la de uno grande, y todos los enjambres presentarían un metabolismo elevado a temperaturas externas relativamente bajas. Las mediciones no apoyan esta suposición. La tasa metabólica era sorprendentemente baja en todos los enjambres, acercándose por término medio al metabolismo de reposo de las abejas, con sólo un ligero aumento a temperaturas del aire más bajas. Además, en enjambres de 1800 a 16.000 abejas, en los que la tasa de enfriamiento pasivo variaba por un factor de ocho, había una pequeña, o nula, tendencia a que la tasa metabólica de los grupos pequeños fuera superior a la de los grandes. De hecho, algunas de las tasas metabólicas más elevadas se registraron en enjambres grandes. Dado que, por razones geométricas, un enjambre grande tiene una mayor proporción de sus abejas en el centro que otro pequeño, y puesto que la tasa metabólica de reposo es una función directa de la temperatura corporal, las tasas metabólicas observadas de enjambres grandes pueden explicarse sobre todo a través de la producción pasiva de calor por parte de abejas acaloradas y no en función de la producción de calor para mantenerse calientes.

Las mediciones en bruto no revelaban cuánto esfuerzo habían de poner las abejas en la generación activa de calor, pero cálculos basados en la tasa de enfriamiento pasivo de las abejas permitieron hacer una estima. Una abeja situada en el núcleo y que se enfriara sólo pasivamente precisaría consumir oxígeno a una tasa de 0,39 mililitros por gramo de peso corporal y por hora para mantener su temperatura. Esta tasa calculada es unas 10 veces inferior al consumo que inevitablemente precisaría el metabolismo inactivo de una abeja en reposo con una temperatura corporal de 35 grados C. De ahí







TEMPERATURAS INTERNAS de un enjambre sometido a una temperatura externa baja, de 2,5 grados Celsius, registrados en un enjambre cautivo. Cada isoterma representa una diferencia de temperatura de un grado C. La primera lectura (*a*) corresponde a un enjambre de unas 5300 abejas, que contenía una reina; el enjambre se dividió entonces en dos, uno de ellos con una reina (*b*) y el otro sin ella (*c*).

que, aun cuando la temperatura en el exterior de un enjambre sea de cinco gélidos grados, las abejas situadas en el centro de un enjambre grande tendrían problemas para librarse del exceso de calor. Cálculos similares indican que en un enjambre pequeño (de unas 1000 abejas) los individuos del núcleo vendrían a requerir las mismas tasas de producción pasiva de calor y pérdida pasiva de calor para mantener su temperatura central cerca de los 35 grados. En un enjambre todavía más pequeño, hasta las abejas del núcleo tendrían que tiritar para mantener una temperatura de 36 grados.

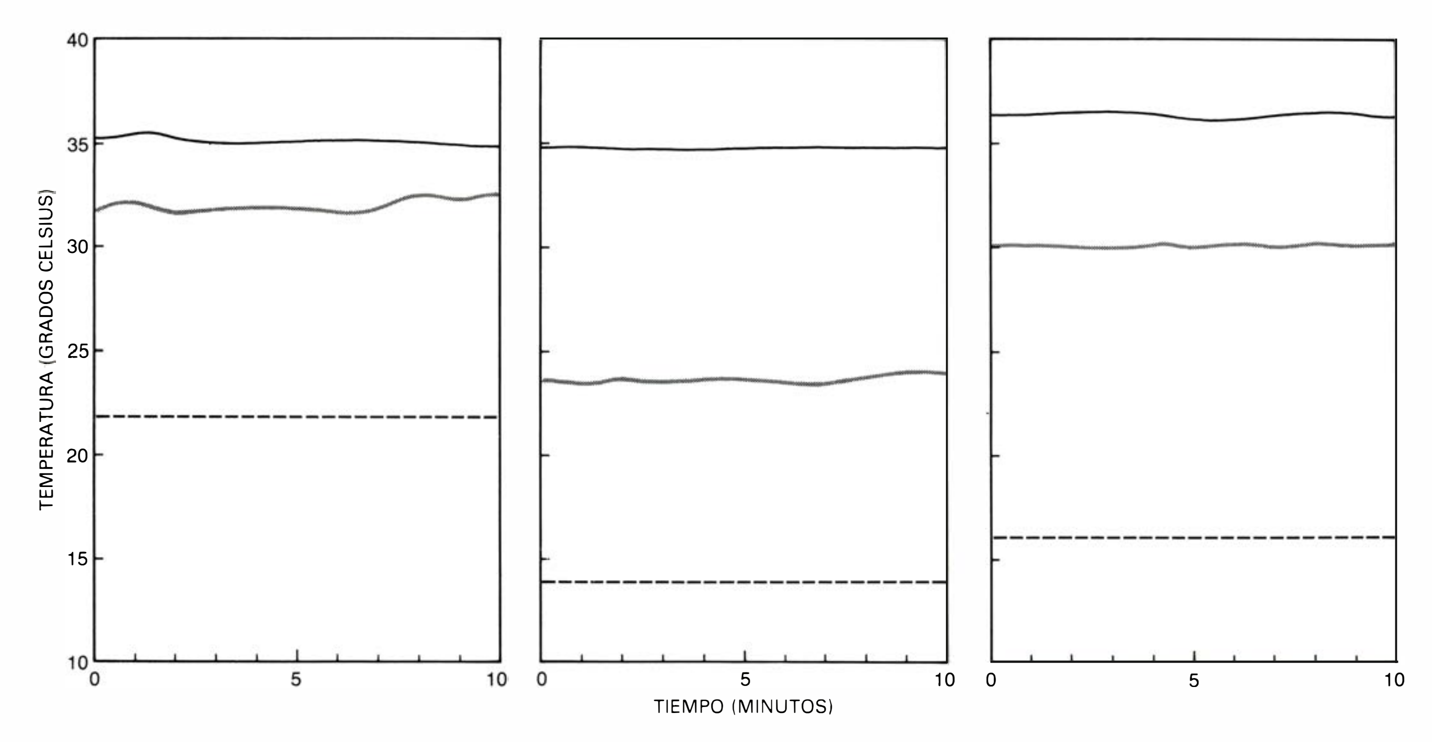
Los datos apuntaban hacia lo siguiente: la mayoría de enjambres compuesos de 5000 a 30.000 abejas no tendrían problema alguno para caldearse internamente. Puesto que en tales enjambres la temperatura del núcleo se mantiene por lo general alrededor de los 35 grados C, esta temperatura debe representar un punto superior fijo que las abejas evitan sobrepasar de manera activa. Podría así encontrarse un enjambre pequeño, de menos de 1000 abejas, cuya temperatura central fuera algo inferior, y eso es lo que mostraban mis mediciones.

¿Cómo evitan sobrecalentarse las abejas del núcleo? De nuevo el problema puede analizarse examinando primero la respuesta de todo el enjambre y observando luego el comportamiento de los individuos por separado. Cuando la temperatura ambiente en torno de un enjambre sube, éste aumenta mucho en longitud y en diámetro. Al extender de este modo su superficie, consigue una mayor pérdida pasiva de calor. Las abejas del manto se hallan más separadas unas de otras, y su cabeza se dirige hacia afuera. A una temperatura inferior, se agrupan más apretadamente y colocan la cabeza bajo el abdomen de las abejas que se encuentran sobre ellas, como ripias. La disposición con temperaturas altas debe facilitar el flujo de aire y de calor a través del manto.

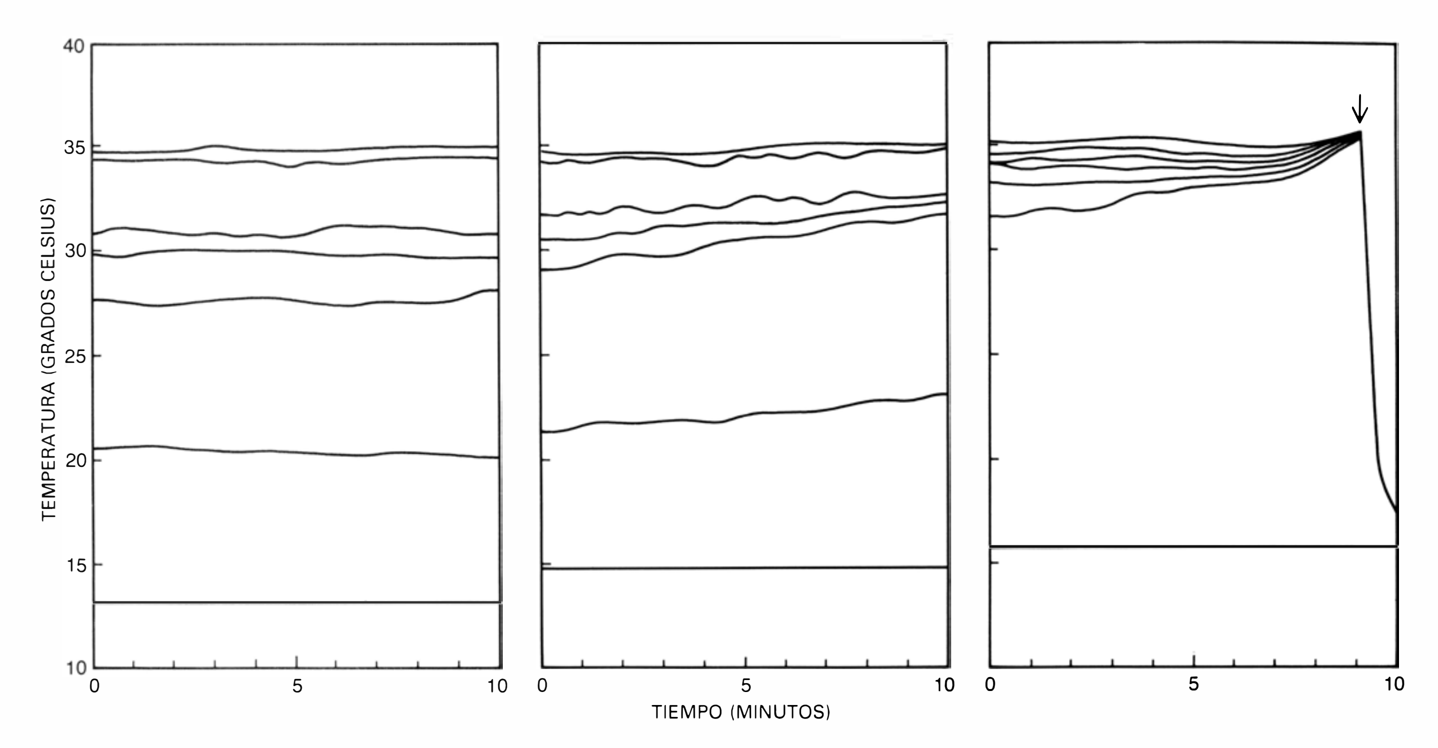
Sin embargo, el exceso de calor se da principalmente en el interior del enjambre. Debe existir un mecanismo que desplace el calor desde el núcleo hasta el manto. Observando enjambres en un contenedor transparente pude apreciar que, a una temperatura ambiente baja, el enjambre era una masa de abejas casi compacta. Al aumentar la temperatura y expandirse el enjambre, se apreciaban cortinas colgantes de abejas inmóviles. Las cortinas formaban pasadizos por los que se desplazaban rápidamente las abejas, desde el centro al manto y, de allí, al centro. Los pasadizos servían asimismo de conductos de ventilación.

La pauta global de la termorregulación en el enjambre no revela de qué modo se coordinan las respuestas. ¿Existe algún tipo de dirección central? ¿Comunican las abejas del núcleo, mediante señales sonoras o químicas, a las abejas del manto que controlan la tasa de pérdida de calor del enjambre la necesidad de calentar o de enfriar? ¿Actúa acaso cada abeja de manera independiente?

Ningún experimento consiguió demostrar una comunicación o cooperación directas. En primer lugar, los enjambres con reina y los que carecían de ella mantenían la misma temperatura central. Este hallazgo eliminaba la posibilidad de que señales químicas procedentes de la reina, situada en el interior del enjambre, actuaran como instrucciones de termorregulación. Fabriqué después un enjambre donde las abejas del núcleo estaban separadas de las del manto por una fina gasa. El enjambre presentaba el mismo perfil de temperatura que los demás enjambres, lo que ponía de manifiesto que, aun cuando las abejas del centro no podían comprobar individualmente la temperatura que rodeaba el enjambre, el núcleo se mantenía a una temperatura constante, con independencia de las fluctuaciones externas. En tercer lugar, cuando registré en cinta magnetofónica los sonidos generados por las abejas del núcleo o del manto a temperaturas externas bajas y reproduje los sonidos en el núcleo o en el manto en momentos de temperatura externa alta (y viceversa), no se provocó ningún cambio en la temperatura del núcleo. Finalmente, el



REGISTROS DE TEMPERATURA de un enjambre libre de 5000 abejas, realizados en continuo a lo largo de un período de varios días y que se reflejan aquí en tres momentos: las 6.30 de la tarde del 7 de junio; las 3 de la madrugada del 8 de junio y las 10 de la mañana del mismo día. Las líneas negras discontinuas indican la temperatura ambiente. La gris registra la temperatura del manto (perímetro) del enjambre y, la negra continua, la del núcleo.



TEMPERATURA DE UN ENJAMBRE antes de emprender el vuelo para desplazarse a un nuevo nido, registrada a las 7, a las 8 y a las 10 de la mañana del día 4 de junio. La flecha señala el momento en que el enjambre echó a volar. La curva inferior registra la temperatura ambiente; la siguiente, la temperatura del manto; las otras tres curvas, la temperatura en puntos situados aproximadamente a medio camino entre el manto y el centro del enjambre, y las dos curvas superiores señalan la temperatura central. Las temperaturas convergieron hacia la del centro en las dos horas anteriores al despegue.

bombeo de aire desde el núcleo de un enjambre mantenido a cinco grados C hasta otro a 20 grados (y viceversa) no incidía en la temperatura de ninguno de los dos centros. Los resultados indican que en la respuesta termorreguladora del enjambre no se hallan implicados sonidos ni señales químicas. Aparentemente, las abejas del núcleo del enjambre no se "cuentan" mutuamente la temperatura local.

Resultaba más provechoso examinar el comportamiento de cada una de las abejas por separado. El biólogo alemán Anton Büdel había ya observado que las abejas del manto tendían a ser, de forma predominante, las obreras viejas, las forrajeadoras, y que las jóvenes abejas "caseras" tendían a encontrarse en el interior del enjambre. Además, M. Delia Allen, del Instituto de Agricultura del Norte de Escocia, había demostrado que las abejas viejas regulaban mejor la temperatura del cuerpo, en especial mediante el estremecimiento. Las abejas jóvenes tienen una capacidad limitada de tiritar y de generar calor propio. En experimentos en los que se ofrece a las abejas la posibilidad de elegir entre varias temperaturas, las jóvenes escogen temperaturas más elevadas que las viejas. Si, en un enjambre, las abejas se comportan de manera similar, estos hechos encajan con mi modelo de regulación de temperatura, en el que las abejas del núcleo se calientan pasivamente y las abejas del manto tiritan para generar calor con el fin de evitar enfriarse por debajo de los 13 grados C, la mínima temperatura corporal por debajo de la cual pierden coordinación motriz y caen del enjambre.

Implanté en unas abejas un termopar piliforme embutido en el tórax y las devolví al manto. Las mediciones de temperatura mostraron que regulaban su temperatura torácica, manteniéndola por encima de los 13 grados. Si se las molestaba, se arrastraban hacia el interior del grupo y su temperatura torácica aumentaba, en cuestión de minutos, a más de 30 grados. De este modo supe que las abejas del manto funcionaban a temperaturas muy por debajo de la del núcleo, y, también, que obtenían considerable calor de sus compañeras, tanto en el manto como en el centro.

Es evidente que, sin comunicación directa, las abejas del manto afectan a (y son afectadas por) las respuestas termorreguladoras de las abejas del interior. Si la temperatura del núcleo es alta, las abejas del manto reciben calor del núcleo, que reemplaza al calor que pierden al ambiente; se retarda así su tasa de enfriamiento y se reduce su carga de producción de calor para la termorregulación. En efecto, puesto que el exceso de calor que debe disiparse se genera pasivamente en el interior del enjambre, ayuda a la termorregulación del núcleo que las abejas del manto no generen calor, puesto que el gradiente de temperatura en el enjambre va hacia afuera, facilitándose el flujo de calor en ese sentido.

¿Tiritan las abejas del manto para mantener una elevada temperatura en el interior o pretenden simplemente regular su propia temperatura a unos 19 grados C? Los datos concuerdan con la hipótesis de que sólo regulan su propia temperatura, aunque, al hacerlo, ayudan también a las abejas del núcleo. Cuando se enfrían las abejas del manto en una temperatura ambiente baja, se apretujan en el enjambre, taponando los pasadizos para evitar la pérdida convectiva de calor desde el centro y reduciendo a la vez la superficie y la porosidad del manto. Comprobé la hipótesis transfiriendo un enjambre desde un ambiente de aproximadamente un grado C hasta la temperatura del laboratorio (de unos 20 grados). Como la hipótesis preveía, la respuesta inmediata fue un descenso en la temperatura del núcleo y un aumento en la temperatura del manto, al aflojar las abejas del manto su apretada formación.

Alternativamente, al hacinarse y tiritar las abejas del manto retardan la pérdida de calor desde el núcleo. Si han ido muy lejos en la retención del calor del enjambre, las abejas del núcleo pueden contrarrestar en parte el efecto

5°
19°
25°
30°
35°

30°
31°
35°

ESQUEMA DE LA REGULACION TERMICA de un típico enjambre de abejas, indicado para una temperatura ambiente baja (*izquierda*) y para una alta (*derecha*). A baja temperatura, las abejas se apretujan, las del perímetro adoptan una formación como de ripia, las abejas de la capa adyacente tiritan para generar calor y las abejas situadas cerca del centro emiten calor a su nivel de metabolismo de reposo y forman corredores para disipar el calor. Cuando la temperatura externa es alta, las abejas del manto adoptan una posición más abierta, la mayoría de abejas se encuentran en un metabolismo de reposo y el enjambre forma varios corredores que contribuyen a la disipación del calor (*flechas*). Así, el núcleo se mantiene a la misma temperatura.





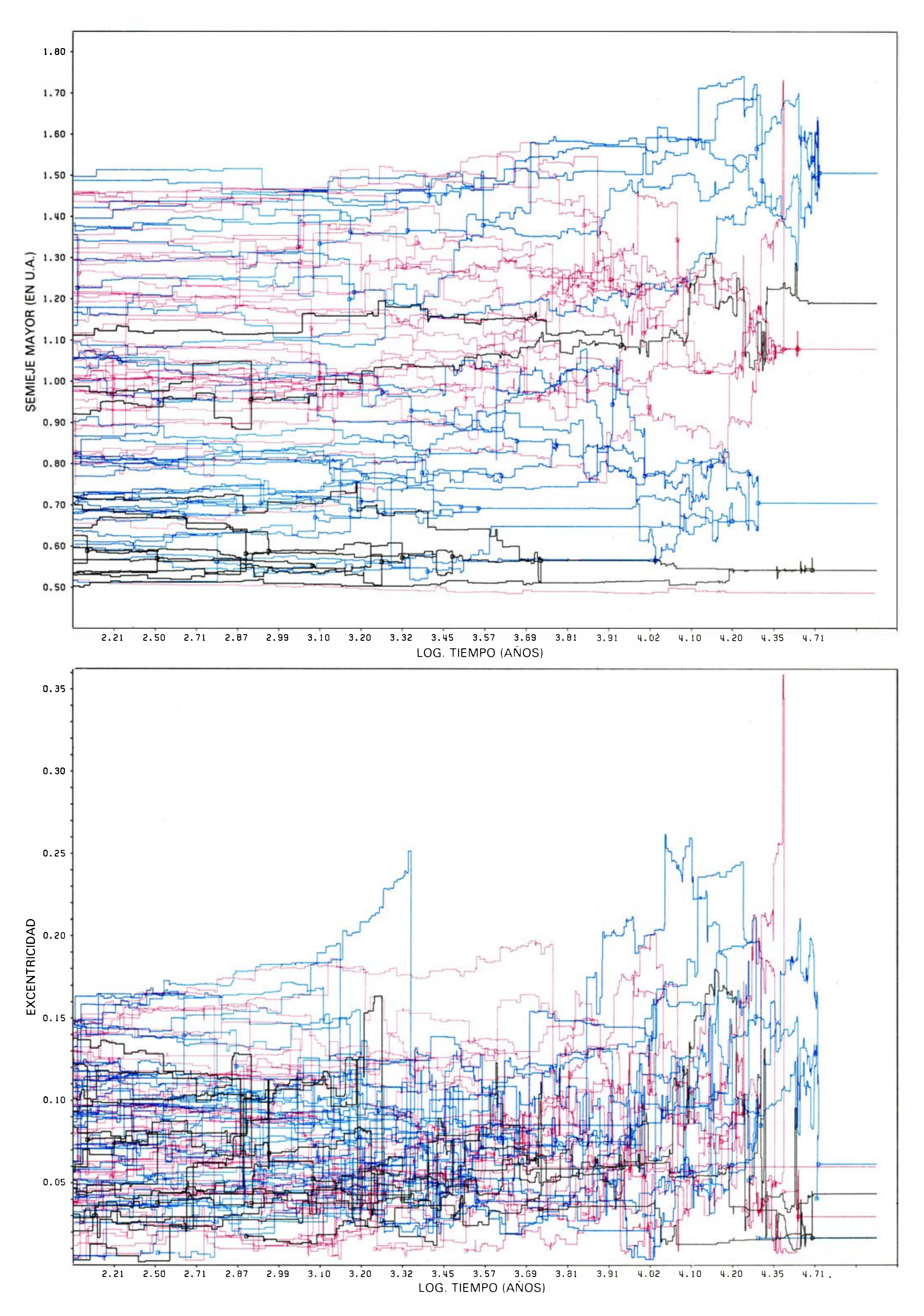
ABEJAS DEL MANTO, que adoptan posiciones distintas según la temperatura externa sea baja o alta. Cuando la temperatura exterior es de tres grados C (*izquierda*), las abejas están muy apretadas; la mayoría de las abejas colocan su cabeza debajo del abdomen de las que se encuentran por encima. A una temperatura externa de 25 grados Celsius (*derecha*), las abejas del manto están más espaciadas y la mayoría presentan la cabeza orientada hacia afuera.

desplazándose hacia el exterior, con lo que crean corrientes de convección. No hay pruebas de que las abejas de una parte del enjambre sepan qué temperatura hay en otra parte del mismo y, en consecuencia, modifiquen su comportamiento. Las abejas actúan independientemente, pero el resultado es una respuesta coordinada que beneficia a todo el enjambre.

Es importante que todas las abejas puedan marchar juntas a ocupar un emplazamiento adecuado para establecer el nido, una vez descubierto y sopesado. La termorregulación del enjambre es crítica, en todos los aspectos, a la hora de buscar un nuevo nido y ocuparlo. En todo momento, el manto ha de mantenerse por encima de los 15 grados C, de lo contrario no podría incitarse a las abejas del mismo a que emprendieran el vuelo.

Ya no resulta tan evidente la razón por la que el núcleo debe mantenerse a unos 35 grados C, en particular por la noche. Sin embargo, esta temperatura es la que requieren para remontar rápidamente el vuelo. Una abeja melífera precisa de cuatro a seis minutos para calentarse, tiritando, de unos 20 grados a los 35 grados. No representa por tanto una gran desventaja dejar que la temperatura del manto descienda por debajo de la temperatura de vuelo, siempre que el núcleo se mantenga caliente y las abejas puedan penetrar en él (o tiritar) para calentarse. Recurriendo al metabolismo activo (la acción de tiritar) sólo para generar calor, las abejas del manto prolongan las limitadas reservas de alimento del enjambre.

Las abejas melíferas que están a punto de enjambrar se atiborran de miel antes de abandonar su colmena, que les servirá, además de provisión de combustible para la termorregulación en espera de ocupar una nueva colmena, de combustible especial para las exploradoras. También es el sustrato a partir del cual se construyen los panales iniciales de la nueva colmena. Sin embargo, las abejas no pueden recurrir con demasiada frecuencia al gasto de energía para regular la temperatura del enjambre, pues la respuesa de éste ante el descubrimiento de un nuevo domicilio sería demasiado lenta. La termorregulación permite una respuesta rápida en una variedad de condiciones meteorológicas y en el estado más crítico del ciclo de la colonia: la competencia por un emplazamiento adecuado para el nido, que posteriormente servirá de manera indefinida para la nueva colonia.



1.80
1.70
1.60
1.50
1.40
1.30
1.20
1.10
1.00
0.90
0.80
0.70
0.60
0.50
SEMIEJE MAYOR (EN U.A.)
2.21 2.50 2.71 2.87 2.99 3.10 3.20 3.32 3.45 3.57 3.69 3.81 3.91 4.02 4.10 4.20 4.35 4.71
LOG. TIEMPO (AÑOS)
0.35
0.30
0.25
0.20
0.15
0.10
0.05
EXCENTRICIDAD
2.21 2.50 2.71 2.87 2.99 3.10 3.20 3.32 3.45 3.57 3.69 3.81 3.91 4.02 4.10 4.20 4.35 4.71
LOG. TIEMPO (AÑOS)

Formación de la Tierra a partir de planetesimales

Los planetas interiores del sistema solar pudieron formarse por coalescencia de un enjambre de pequeños cuerpos en colisión, en órbita alrededor del Sol. El proceso puede simularse por ordenador

George W. Wetherill

Una explicación cabal de cómo se formó la Tierra requiere conocer primero cómo se condensó el Sol a partir de una nube de gas y polvo, y cómo una pequeña proporción de materia de esta nube dejó de caer en el Sol para formar, por su parte, los planetas. Pero no es fácil adquirir ese conocimiento. Por un lado, las simulaciones realizadas con ayuda de ordenador sugieren que si semejante nube o nebulosa solar era lo suficientemente densa como para sufrir un colapso gravitatorio, generaría un número considerable de sistemas múltiples, en lugar de una estrella simple. Y en caso de que tratáramos de explicar la condensación de un solo sol, ello no facilitaría la descripción del origen de los "planetas terrestres" (los cuerpos rocosos Mercurio, Venus, Tierra y Marte). La masa de los planetas terrestres es sólo de un 0,0005 por ciento de la solar. No es, pues, lógico esperar que un detallado conocimiento de un proceso tan diminuto como la formación de la Tierra pueda deducirse de una sencilla consideración de las propiedades generales de la nebulosa solar.

Sin embargo, hay vías más prometedoras de abordar el asunto. A pesar de la complejidad del tema referente a la formación del sistema solar a partir de una nebulosa, existen sólo dos procesos básicos capaces de explicar la formación de un planeta terrestre. Este pudo surgir de una concentración local de gas y polvo nebular, que habrían colapsado hasta constituir un planeta por su propia gravitación, o bien el planeta pudo haberse constituido por acumulación de pequeños cuerpos (planetesimales), previamente condensados a partir del gas y el polvo. A través de la simulación de estos dos procesos alternativos, los teóricos han ensayado una interpretación de cómo llegó a existir el sistema solar. Por ejemplo, es posible especificar qué condiciones dadas en la nebulosa solar hicieron viable la formación de los planetas, y distinguirlas de aquellas otras que la hubieran hecho inviable. También es posible determinar lo que una y otra hipótesis prevén acerca de determinadas propiedades, tales como la temperatura inicial de la Tierra. Con un poco de optimismo podemos esperar que una comparación de estas predicciones con los registros geológicos nos muestre cuál alternativa representa el proceso más probable de formación de los planetas terrestres.

En este artículo vamos a describir la interpretación actual sobre la formación de la Tierra a base de planetesimales. Es demasiado pronto para decidir si la Tierra se formó de este modo, y no por la alternativa de colapso gravitacional. Con todo, resulta perfectamente verosímil que los planetas terrestres se hayan formado a partir de planetesimales. Si ello fue así, las simulaciones muestran que su formación estuvo virtualmente completa alrededor de 100 millones de años después de haber comenzado, y que la Tierra recién formada estaba tan caliente, que la mayor parte de su masa interna se encontraba fundida.

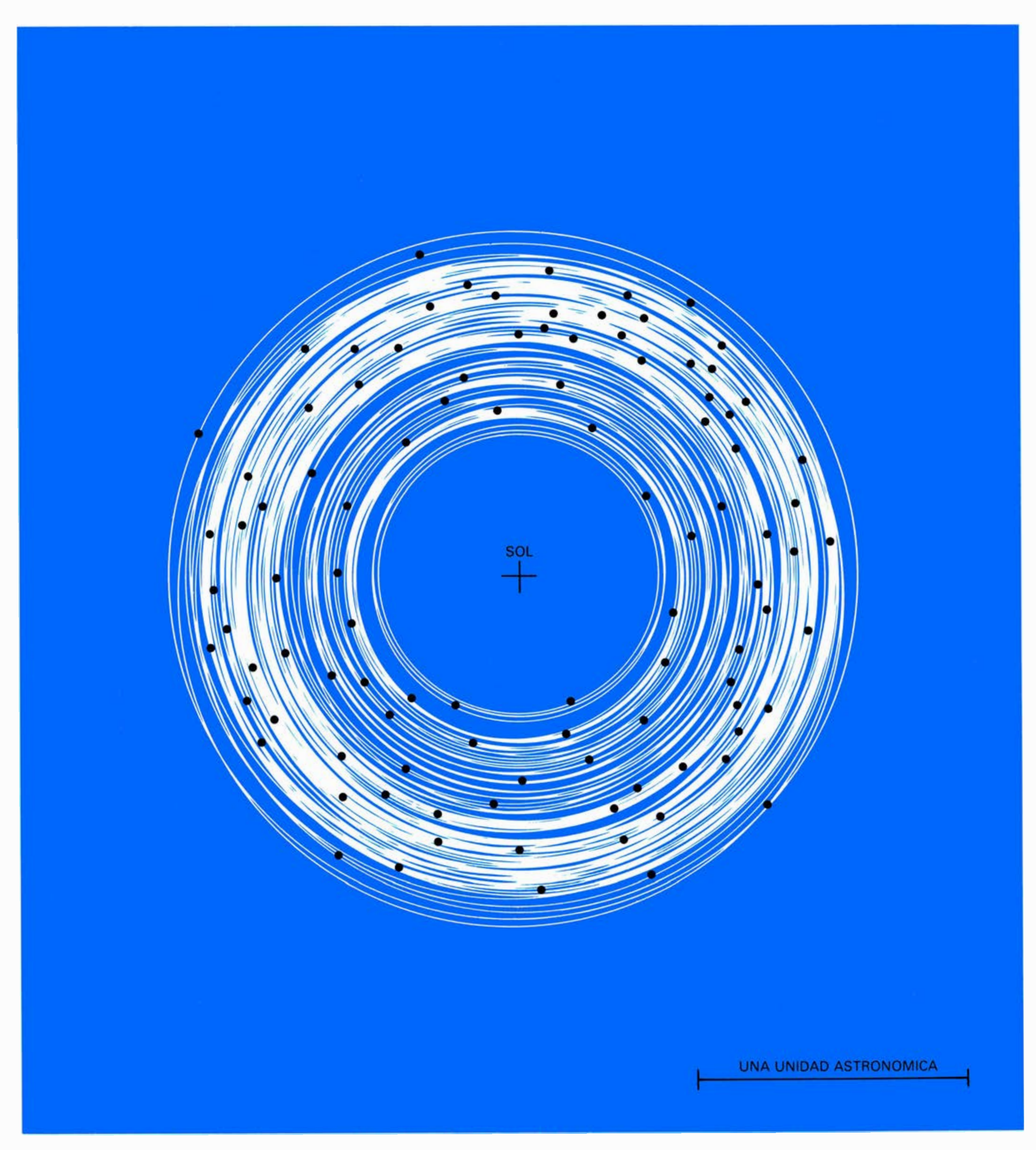
La primera cuestión que hay que plantearse en la "construcción" de los planetas a partir de planetesimales es si, cuando dos cuerpos colisionan, se agregan por coalescencia para formar un cuerpo mayor. Es obvio que una colisión no siempre se traduce en coalescencia. Por ejemplo, si dos rocas de un metro de diámetro chocan entre sí con una velocidad relativa de 1000 kilómetros por hora (0,36 kilómetros por segundc), se partirán en fragmentos que saldrán despedidos en todas direcciones. Por este camino no pudo llegarse a la formación de los planetas.

Una colisión a una velocidad relativamente baja pudo, en cambio, permitir la coalescencia. ¿Cuál tuvo que ser esa velocidad? A la hora de contestar esta pregunta, hay que tener especialmente en cuenta la magnitud llamada velocidad de escape gravitatorio. Es la

ACRECION DE LOS PLANETAS del sistema solar interno a partir de la colisión de pequeños cuerpos llamados planetesimales, según la procesó, por ordenador, Larry P. Cox, del Instituto de Tecnología de Massachusetts. La simulación comienza considerando 100 cuerpos de igual masa. De acuerdo con las leyes de Kepler, cada cuerpo seguía una órbita elíptica alrededor del Sol. Cada órbita quedaba especificada por dos magnitudes: su semieje mayor (*gráfico parte superior*) y su excentricidad (*gráfico parte inferior*). El semieje mayor es la medida de la distancia de la órbita al Sol; está indicada en unidades astronómicas, o U.A. (Una U.A. equivale a la distancia media entre la Tierra y el Sol, que es, aproximadamente, de $1,5 \times 10^8$ kilómetros.) La excentricidad mide el grado de separación de una elipse respecto de la circularidad. Una excentricidad 0 significa una circunferencia; una excentricidad 1 significa una parábola, en la cual el semieje mayor se prolonga indefinidamente. Aunque los parámetros iniciales de las órbitas se generaron al azar, se ha escogido como valor máximo una excentricidad inicial de 0,15, y los semiejes mayores se han supuesto que abarcan un ámbito entre 0,5 y 1,5 U.A. La simulación es bidimensional: considera a todas las órbitas en un plano. Las colisiones entre los cuerpos en un plano son más probables que en un espacio tridimensional; de aquí que la acreción en esta simulación "plana" se complete en sólo 61.000 años. Los pequeños cuadraditos en los gráficos indican colisiones, que hacen que los cuerpos que chocan se unan. Las discontinuidades en algunos cuadraditos del gráfico inferior indican grandes descensos en excentricidad. Al final, quedan solamente seis planetas. Los diferentes colores distinguen las secuencias mediante las cuales se fueron formando. El planeta sobreviviente más cercano al Sol es un planetesimal que nunca colisionó, aunque las variaciones en su trayectoria muestran que su órbita se vio repetidamente perturbada por acercamientos de otro cuerpo. Los otros cinco planetas son mucho mayores. Podemos llamarlos, si se nos antoja, Mercurio, Venus, Tierra, Luna y Marte. El éxito de la simulación estriba en que el resultado final recuerda en alto grado a los planetas terrestres, con sus masas y sus órbitas. Hemos distorsionado el eje que representa el tiempo al objeto de que iguales números de sucesos perturbadores (colisiones y acercamientos) ocupen iguales intervalos en el grabado.

mínima velocidad relativa necesaria para que dos cuerpos en movimiento puedan liberarse de su mutua acción gravitatoria. Para escapar de la Tierra, por ejemplo, un proyectil lanzado desde su superficie debe alcanzar una velocidad de alrededor de 11 kilómetros por segundo. Para escapar de la Luna, es suficiente una velocidad de 2,4 kilómetros por segundo. Si dos planetesimales llegaran a colisionar precisamente a su velocidad de escape, y no gastaran energía, tendrían exactamente la energía necesaria para rebotar y recobrar trayectorias independientes.

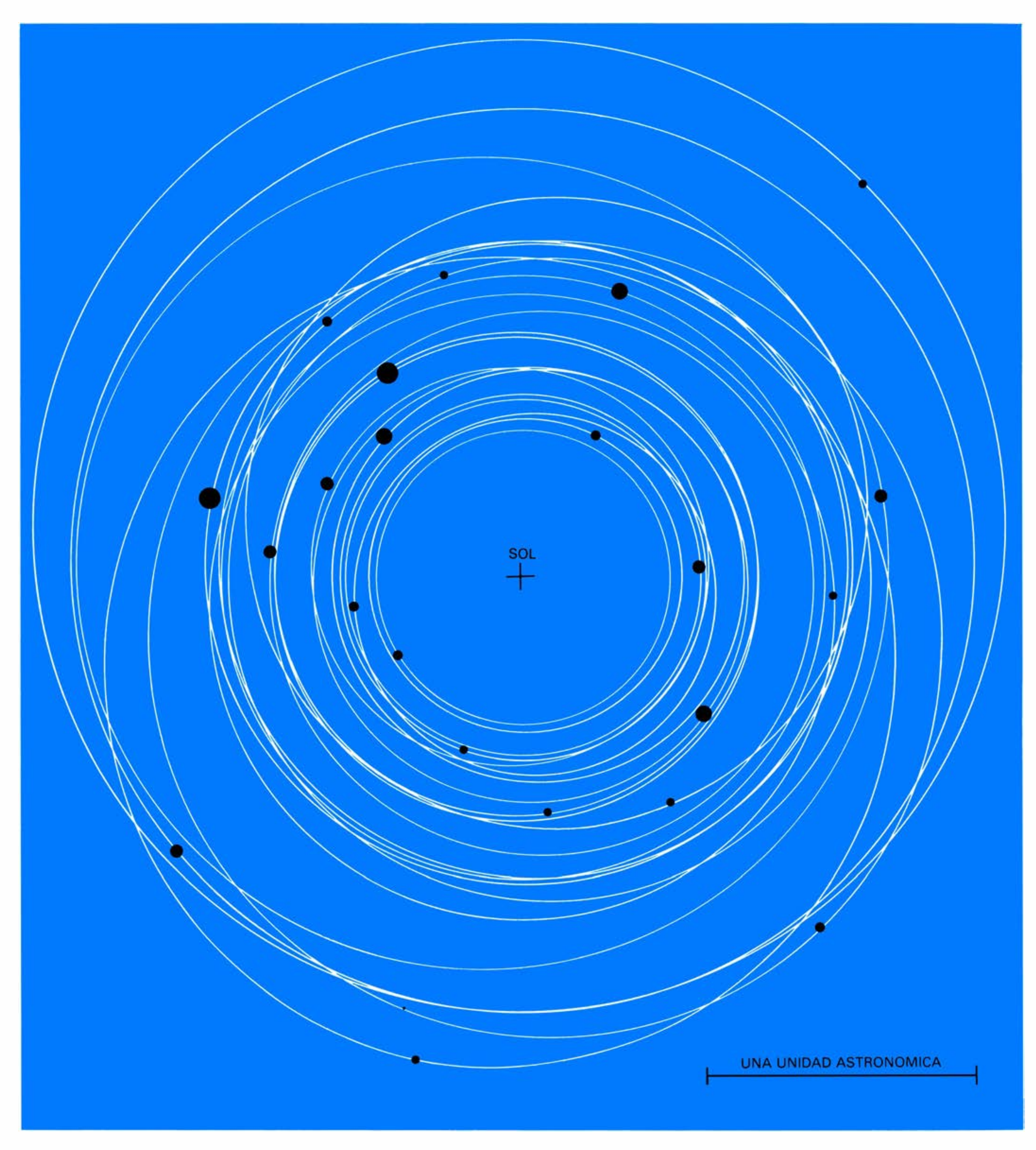
En una colisión real, los plantesimales tienen que perder energía, porque los cuerpos que chocan se deforman o desgarran en una parte de su contenido. Esta energía perdida debe disiparse en forma de calor. Los dos planetesimales que colisionan exactamente a su velocidad de escape se ven así incapaces de huir el uno del otro, y sus frag-



SIMULACION TRIDIMENSIONAL de la formación del sistema solar interno, iniciada también con 100 planetesimales. A cada uno se le ha asignado una masa de 1,2 × 10^{26} gramos, de modo que la masa total (1,2 × 10^{28} gramos) sea la de los rocosos planetas internos, desde Mercurio hasta Marte, junto con sus satélites. De nuevo, los parámetros iniciales de las órbitas se han determinado al azar. La ilustración muestra las elipses iniciales, pero no sugiere su inclinación mutua. Como se supone que los planetesimales se formaron en una delgada lámina de polvo en el plano central de la nube que constituyó el sistema solar, las inclinaciones no deben sobrepasar los cinco grados. En un choque real, los planetesimales que colisionan tienen que perder energía.

mentos deben combinarse para formar un solo cuerpo mayor. Realmente, los experimentos efectuados por William K. Hartmann en el Instituto de Ciencia Planetaria de Tucson, Arizona, y los cálculos numéricos realizados por Thomas J. Ahrens y John D. O'Keefe, del Instituto de Tecnología de California, han mostrado que, cuando la velocidad de impacto es de dos veces, y en ocasiones hasta casi tres veces la velocidad de escape, los planetesimales pueden fusionarse.

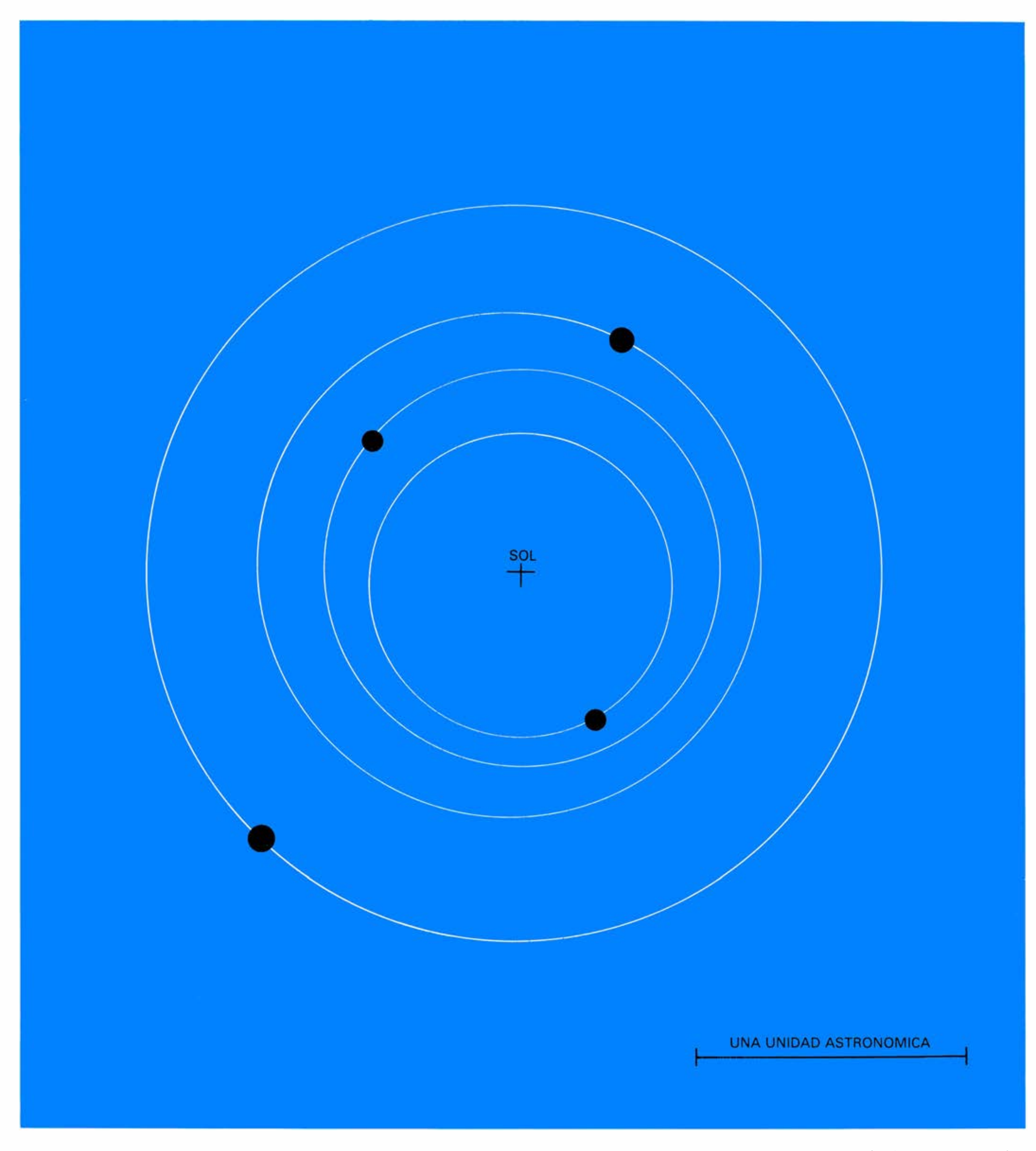
Algunos nuevos razonamientos nos permitirán progresar en esta idea. La velocidad de impacto de dos cuerpos viene determinada por la velocidad relativa que los anima cuando están distantes uno de otro, incrementada por la velocidad que adquieren cuando al acercarse se aceleran mutuamente por efecto de la atracción gravitatoria. Si la velocidad relativa en el momento de la colisión llega a ser el doble de la veloci-



ESTADIO INTERMEDIO de la simulación, que muestra el desenvolvimiento del sistema solar interno, cuando han pasado 30,2 millones de años, y perviven 22 cuerpos. Sus órbitas se han hecho más marcadamente elípticas, y sus velocidades relativas son mayores. Además, su intervalo de distancias al Sol ha aumentado: las órbitas más internas están más cercanas al Sol, y las más externas se han alejado. Las mutuas inclinaciones entre las órbitas han disminuido considerablemente la incidencia de las colisiones, en comparación con la incidencia en que se ha presentado la de los acercamientos. De aquí que hayan de transcurrir decenas de millones de años de tiempo simulado para que el proceso finalice. En una simulación bidimensional basta con 61.000 años.

dad de escape, la velocidad relativa cuando los dos cuerpos se encontraban aún a gran distancia era $\sqrt{3}$ (aproximadamente 1,7) veces la velocidad de escape. Si la velocidad de colisión triplica la de escape, la velocidad relativa cuando los cuerpos estaban distantes era de 2 $\sqrt{2}$ (aproximadamente 2,8) veces la velocidad de escape. Entonces, la regla según la cual los cuerpos sólo pueden fusionarse por coalescencia si la velocidad relativa en el momento del impacto es menor que dos o tres veces la velocidad de escape, tiene también correspondencia con la velocidad relativa de los cuerpos en sus libres trayectorias por el espacio, mucho antes de la colisión.

En algunas simulaciones de la formación de planetas a partir de planetesimales, la velocidad relativa de dos planetesimales cualesquiera se considera un parámetro libre: dando por supuesto que la velocidad puede llegar a ser lo



ESTADIO FINAL de la simulación. Quedan cuatro planetas completamente formados en órbitas casi circulares y aisladas, en torno al Sol. Los planetesimales se han fundido por coalescencia. El planeta mayor, el cuarto a partir del Sol, está constituido por 34 cuerpos originales. Aunque la ilustración muestra la simulación al cabo de 441 millones de años, el proceso está casi completo mucho antes. Hay 11 cuerpos al cabo de 79 millones de años y 6 al cabo de 151 millones de años. Esto sugiere que la Tierra se formó a lo largo de un intervalo de tiempo que osciló alrededor de 100 millones de años. La simulación fue hecha por el autor, sirviéndose de un ordenador en el Departamento de Magnetismo Terrestre de la Institución Carnegie de Washington.

suficientemente baja como para que los planetesimales se agreguen por coalescencia. Por otro lado, la velocidad puede ser también demasiado baja. Las leyes de los movimientos planetarios, formuladas por primera vez por Johannes Kepler, precisan que la órbita que describe un cuerpo alrededor del Sol es una elipse. La velocidad en un punto cualquiera de esa órbita está determinada por las dimensiones y la forma de esa elipse. De ello se sigue que la velocidad relativa de dos cuerpos en órbita alrededor del Sol tenderá a ser grande, a no ser que las órbitas sean similares en amplitud, forma y orientación. En este caso concreto de baja velocidad relativa, la posibilidad de coalescencia y aumento de tamaño se restringe a una pequeña porción de planetesimales, especialmente aquellos cuyas órbitas son muy similares entre sí. La formación de planetas terminará cuando aquellos planetesimales se hayan combinado, y el simulado sistema solar interno recordará más los anillos de Saturno que una serie de cuatro planetas terrestres en medio de un espacio vacío.

La eventualidad de una velocidad relativa cercana a la velocidad de escape constituye, pues, el mejor procedimiento para lograr una simulación acertada. Ella podría facilitar la coalescencia de planetesimales en colisión y al mismo tiempo multiplicaría el número de potenciales parejas colisionantes. Pero se plantea otro problema. Tan pronto como los cuerpos aumentan de tamaño, la velocidad de escape se incrementa también: para un cuerpo de densidad uniforme, la velocidad de escape es proporcional a su radio. Si un planetesimal con un radio de un metro tuviese una velocidad de escape de 15 centímetros por segundo, un planetesimal de un kilómetro de radio debería tener una velocidad de escape de 1,5 metros por segundo, y un planetesimal con un radio de 1000 kilómetros tendría una velocidad de escape de 1,5 kilómetros por segundo. Cada uno de esos cuerpos puede continuar creciendo por acreción si colisiona con un cuerpo que se mueve a una velocidad no mayor de dos o tres veces su velocidad de escape. Así es como la velocidad relativa de los planetesimales puede "guardar el paso" con su crecimiento y con su progresiva velocidad de escape, si esta se mantiene en su grado óptimo.

La posibilidad de que la velocidad relativa de los planetesimales pueda mantenerse naturalmente proporcional a la creciente velocidad de escape fue introducida en 1950 por L. E. Gurevich y A. I. Lebedinskii, del Instituto de Geofísica Aplicada de Moscú. La idea fue desarrollada con mayor amplitud por Víctor S. Safronov, del mismo Instituto, a lo largo de la década de 1960. La obra de Safronov era analítica: desarrollaba expresiones matemáticas para la evaluación de cantidades tales como la velocidad relativa promedio de los miembros de un enjambre de planetesimales. Su modelo más simple describía el enjambre en ausencia de frenado o rémora alguna causada por el gas o la nebulosa solar.

Las velocidades relativas de los cuerpos en el enjambre quedan determinadas por el balance de dos efectos contrarios. Cuando los cuerpos pasan cerca uno de otro, sin colisionar, su atracción gravitacional perturba sus órbitas. La perturbación cambia la velocidad relativa, con lo cual cada cuerpo se encuentra sucesivamente con otros moviéndose en órbitas que se intersecan. A veces, la velocidad relativa después del acercamiento es mayor que la que el cuerpo hubiera mantenido de no haberse producido aquél; otras veces, se hace menor. El resultado neto, sin embargo, es un aumento de la velocidad relativa promedio. Las colisiones producen el efecto contrario: tienden a hacer las órbitas de los planetesimales más circulares de lo que antes eran, y, por consiguiente, más similares. De ello se sigue que las velocidades relativas tienden a decrecer. Así, en un tiempo dado, el incremento de la velocidad como resultado de los acercamientos tiende a ser contrarrestado por la disminución de la velocidad fruto de las colisiones.

Cada uno de estos efectos puede analizarse con mayor detalle. Las leyes newtonianas del movimiento contribuyen a este objeto, juntamente con la teoría del movimiento aleatorio, que describe la naturaleza esencialmente fortuita de las interacciones de los planetesimales. Para pequeñas perturbaciones causadas por acercamientos, el cambio de la velocidad de un cuerpo pequeño (el "proyectil") que va al encuentro de otro mayor (el "blanco") es de $\sqrt{M/D}$, siendo M la masa del blanco y D la distancia en la máxima aproximación. Para acercamientos muy próximos, la distancia D es aproximadamente igual al radio, R, del blanco. Además, la masa M está en función de R^3. La expresión $\sqrt{M/D}$ puede así transformarse en $\sqrt{R^3/R}$, o, simplemente, en R. En pocas palabras, el cambio en la velocidad del proyectil tiene exactamente la misma dependencia directa del radio del blanco que la velocidad de escape. Esta relación implica que el aumento de la velocidad relativa que acompaña a una serie de acercamientos marcha en paralelo al incremento de la velocidad de escape producida por el aumento en tamaño de los cuerpos.

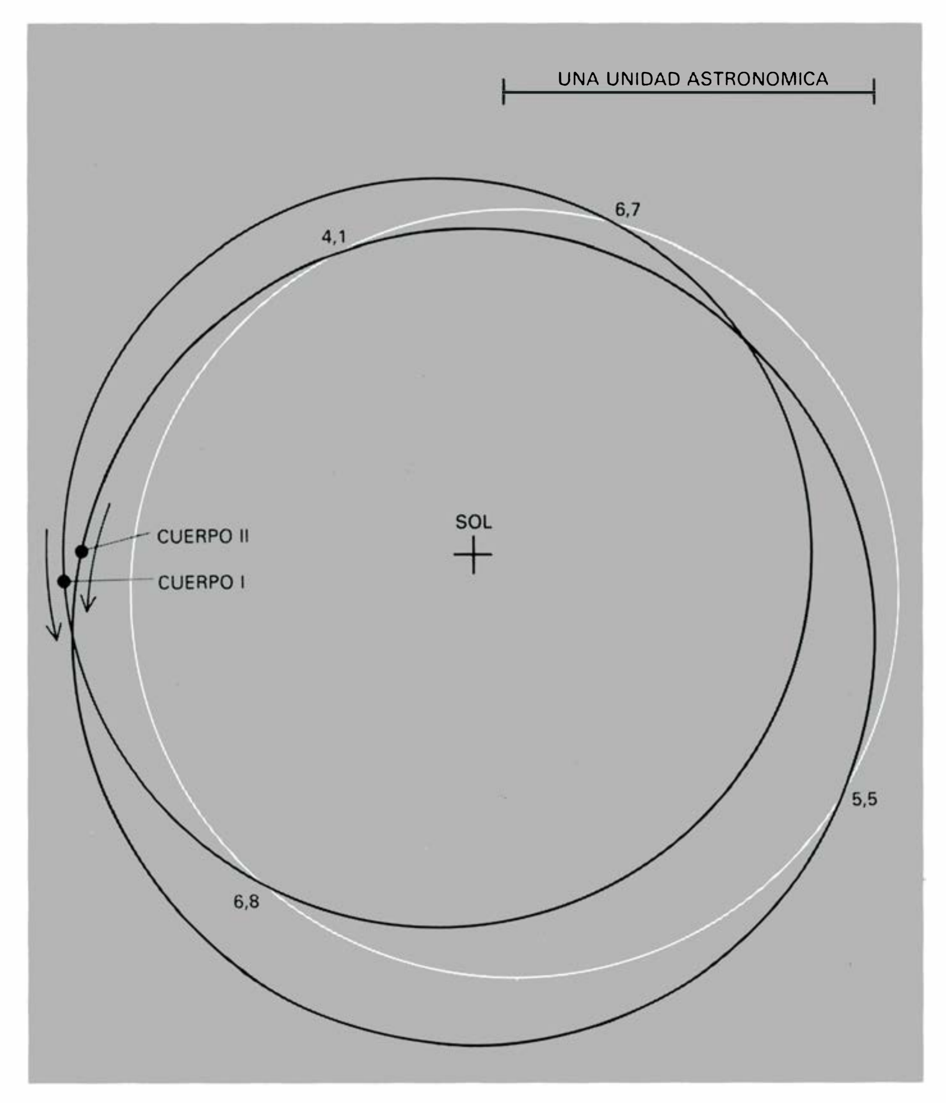
Los acercamientos más distantes también perturban las órbitas. Los acercamientos son significativos a partir de una distancia menor que el radio de lo que pudiéramos llamar esfera de influencia gravitatoria del blanco. Dentro de ese espacio, el movimiento del proyectil está dominado por el campo gravitatorio del blanco. Fuera de él, el movimiento está dominado por el campo gravitatorio del Sol. Si los cambios de velocidad causados por acercamientos distantes están promediados por encima de los que se producen fuera de la esfera de influencia (alrededor de 75 veces el radio del blanco), el resultado es una proporción más o menos fija respecto de los cambios causados por acercamientos a distancias del orden de un radio del blanco. Entonces, los incrementos de velocidad resultantes de los acercamientos distantes también aumentan con la velocidad de escape.

El aumento de la velocidad por acercamientos a distancias comprendidas entre 1 y 75 radios compensa de sobra el efecto ralentizador de las colisiones. Con seguridad, el incremento del área de exposición de cada blanco, conforme éste aumenta de tamaño, aumenta también la frecuencia de las colisiones. Al mismo tiempo, sin embargo, el incremento de la masa del blanco expande su esfera gravitatoria de influencia, y así la probabilidad de una perturbación por acercamiento crece también. El resultado final es curioso: la velocidad relativa en el enjambre de planetesimales se mantiene proporcional a la velocidad de escape. Además, para valores aceptables de la pérdida de energía en colisiones, la velocidad relativa es aproximadamente igual a la velocidad de escape. Esto resulta ser justamente lo que hacía falta para que los planetesimales continúen su proceso de agregación por coalescencia, a medida que aumenta el tamaño de cada cuerpo.

En el arranque del proceso (tal como ha sido establecido por Safronov) los planetesimales son pequeños. Cada uno tiene un radio de pocos kilómetros. La velocidad de escape es también pequeña, y, la velocidad relativa, próxima a la de escape. Solamente los cuerpos con órbitas similares se encuentran y se funden por coalescencia. Ahora bien, la acumulación tiene lugar a lo largo de la región del sistema solar que final-

mente van a ocupar los planetas terrestres. Conforme los cuerpos se engrosan, crece la velocidad de escape. La acumulación continúa, debido a que la velocidad relativa se mantiene cercana a la velocidad de escape. De acuerdo con las leyes de Kepler, el incremento de la velocidad relativa está asociado con el incremento de la excentricidad de las órbitas. En otras palabras, cada órbita tiende a configurar una elipse más pronunciada. Esta tendencia permite a cada planetesimal alcanzar a cuerpos más distantes. Se mantiene así el número de potenciales compañeros de colisión para un determinado planetesimal, a pesar de que el número total de cuerpos está decreciendo y que los cuerpos mismos se encuentran cada vez más separados entre sí.

¿Qué resta por explicar antes de que podamos admitir que la Tierra y los otros planetas terrestres pudieron formarse de esta manera? Podemos preguntarnos cómo llegaron a formarse los pequeños planetesimales que aparecen como punto de partida en el modelo de Safronov. No se trata de una cuestión decisiva. Stuart J. Weidenschilling, del Instituto de Ciencia Planetaria, ha mostrado que las débiles atracciones eléctricas llamadas fuerzas de Van der Waals pueden muy bien haber originado gránulos de polvo en la nebulosa solar, capaces de asociarse unos a otros, hasta formar cuerpos de alrededor de un centímetro de diámetro. Se supone que la nebulosa estaba en rotación. La resistencia al avance de los cuerpos ejercida por el gas en rotación debió haberlos arrastrado hacia el plano central de la nebulosa. (El plano central es perpendicular al eje de rotación de la nebulosa.) El resultado debió haber sido una delgada capa central de polvo. Algunos de los cuerpos se habrían engrosado allí por colisión. Cuando los más voluminosos habrían alcanzado un diámetro de alrededor de un metro, la concentración de materia en la lámina de polvo debió hacerla gravitatoriamente inestable. Los cuerpos debieron ser entonces arrastrados y atraídos, juntamente con el polvo residual, hasta formar planetesimales de alrededor de un kilómetro de diámetro. De ese proceso se han ocupado varios investigadores en los últimos 30 años. Se trata de planetesimales requeridos por la teoría de Safronov. Se necesitarían 10^{13} para constituir los planetas terrestres.



LAS INTERACCIONES ENTRE DOS PLANETESIMALES son los sucesos simulados que hacen que un enjambre de planetesimales llegue a convertirse en el sistema solar interno. Aquí se representan las trayectorias de dos planetesimales que están próximos a intersecarse. Cada cuerpo tiene una masa de 10^{27} gramos y un radio de alrededor de 4000 kilómetros. La órbita de un tercer planetesimal (*en blanco*) representa el movimiento de otro cuerpo del enjambre. La velocidad de cada uno de los cuerpos que interaccionan con respecto a la órbita del tercer cuerpo se expresan en kilómetros por segundo, en los puntos de intersección. Los resultados de la interacción se muestran en las dos ilustraciones siguientes.

Otros problemas son algo más difíciles de abordar. Por ejemplo, el modelo más simple de Safronov describe un enjambre de planetesimales que inicialmente tenían todos la misma masa. En 1962, extendió sus resultados para el caso de cuerpos de diferentes masas; y pudo demostrar que la velocidad relativa iguala la velocidad de escape de los cuerpos más voluminosos solamente si estos poseen la mayor parte de la masa del sistema. En este caso, los cuerpos más pequeños se rompen por colisión, pero los mayores se engrosan. Es posible, sin embargo, que como el sistema contiene un enorme número de pequeños planetesimales, la masa nunca llegue a estar concentrada preferentemente en los cuerpos mayores. El quid de esta cuestión estriba en que hace falta calcular simultáneamente la evolución de la velocidad relativa y la evolución de la distribución de masas. La teoría de Safronov no es adecuada para realizar este cálculo.

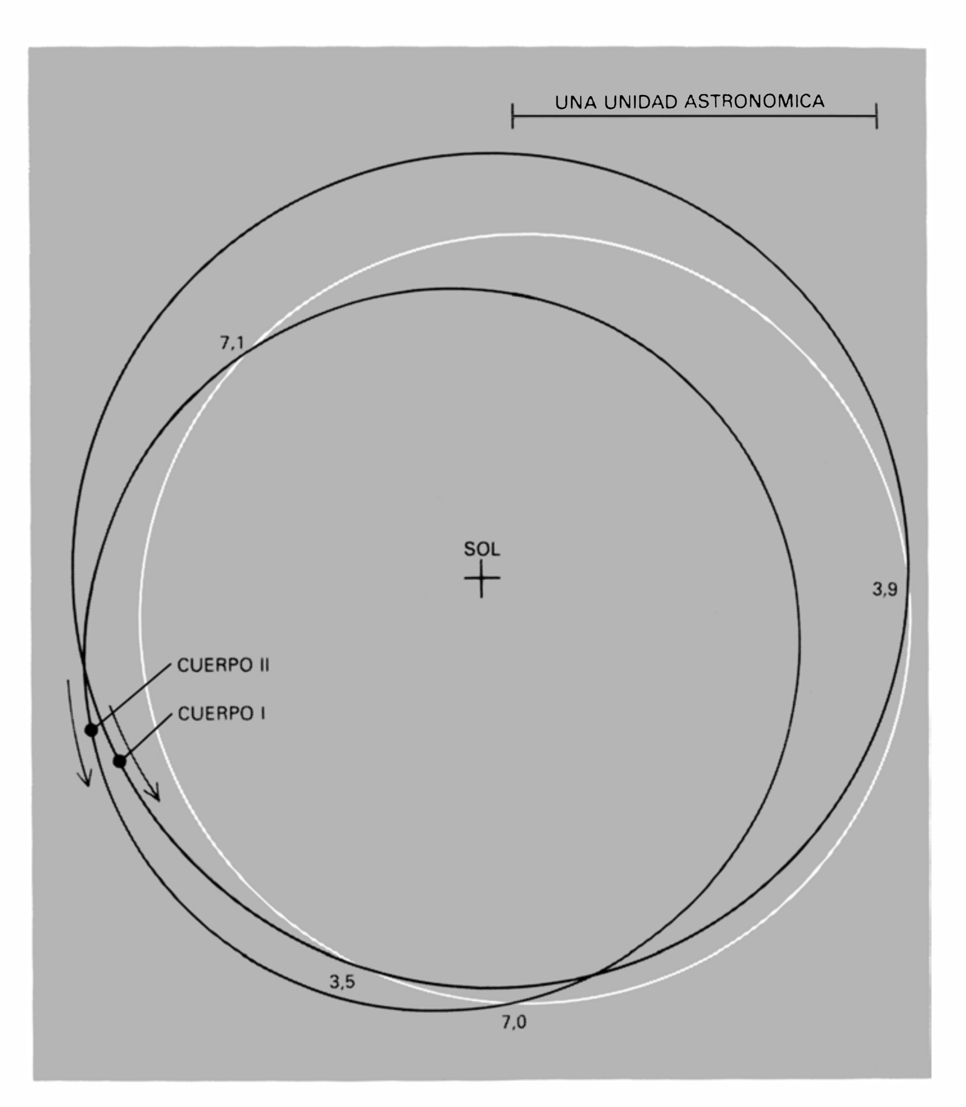
Richard J. Greenberg y sus colegas del Instituto de Ciencia Planetaria han realizado importantes estimaciones a este respecto. Su trabajo está limitado a los primeros estadios de crecimiento en un enjambre de planetesimales, comenzando en el momento en que la mayor parte de la masa del enjambre está concentrada en planetesimales de alrededor de un kilómetro de diámetro. La velocidad relativa de estos cuerpos es sólo de unos pocos metros por segundo, valor comparable al de la velocidad de escape. De acuerdo con los cálculos de Greenberg, algunos de estos cuerpos colisionaron entre sí, y se unieron por coalescencia; y después de unos 15.000 años, algunos se habrían engrosado hasta medir unos 100 kilómetros. La mayor parte de la masa, sin embargo, se encontraba todavía en los planetesimales pequeños. Tal como predice el resultado obtenido en 1962

por Safronov, hay un número demasiado escaso de grandes cuerpos para aumentar la velocidad relativa y superar la ralentización por colisiones.

Es este un momento crucial del proceso. Para cuerpos de un diámetro del orden de ocho kilómetros, resulta que la velocidad relativa se incrementa. Mientras estas bajas velocidades iniciales persisten, los primeros planetesimales que se han engrosado hasta un diámetro de 100 kilómetros llegan a tener una insuperable ventaja sobre sus compañeros más pequeños. En particular, un acercamiento entre un pequeño planetesimal y otro grande, a una velocidad bastante por debajo de la velocidad de escape, puede perturbar la trayectoria del planetesimal pequeño en tal grado, que éste colisiona con el grande, cuando no parecía verosímil que, con su trayectoria inicial, pudiera operarse el choque. Este efecto se conoce por enfoque gravitacional. En relación con la predicción de colisiones, ello incrementa el área seccional del cuerpo grande por encima de su normal relación geométrica, del cuadrado del radio, hasta una dependencia próxima a la cuarta potencia del radio.

Si el enfoque gravitacional es suficientemente eficaz, los primeros cuerpos grandes que se forman capturan a los pequeños, y con ello impiden el crecimiento de cuerpos de tamaño intermedio. La distribución de tamaños entre los planetesimales muestra así dos máximos, representando uno los cuerpos grandes y, el otro, los pequeños que no han sido aún absorbidos. Los cuerpos grandes se mueven en órbitas casi circulares, pues nada ha sucedido que pudiera aumentar la excentricidad de sus órbitas. En efecto, los impactos de los pequeños planetesimales contra los grandes tienden a hacer las órbitas de éstos más circulares.

Si este proceso continúa, los cuerpos grandes acaban adquiriendo la mayor parte de la masa del sistema. Pero entonces puede ser demasiado tarde para que se forme un sistema como nuestro sistema solar. Los cuerpos grandes se mueven por órbitas que no se cortan entre sí, en las cuales acaban capturando a los últimos planetesimales pequeños. El resulado final es un sistema de alrededor de 5000 planetas menores, del orden de 1000 kilómetros de diámetro cada uno. Es concebible que este sistema pudiera tener una segunda oportunidad para formar una serie de planetas terrestres. Tal vez las interacciones gravitatorias de largo radio de miles de planetas menores podrían incrementar la excentricidad de las órbitas lo suficiente como para reanudar el proceso de crecimiento. Pero no es ésta la evolución previsible más correcta en el modelo teórico de Safronov.



ACERCAMIENTO DE DOS PLANETESIMALES con la consiguiente perturbación de sus órbitas, de modo que cada uno sigue una nueva trayectoria elíptica alrededor del Sol. La velocidad del cuerpo I con respecto al cuerpo que no toma parte en la interacción ha decrecido; la del cuerpo II se ha incrementado. (Las velocidades con respecto al tercer cuerpo son importantes, porque los dos cuerpos que se han influido mutuamente son los que menos probabilidades tienen de volver a encontrarse.) Por término medio, los acercamientos aumentan las velocidades relativas de los planetesimales. Este acercamiento se produjo, en el momento de mayor proximidad, a una distancia de tres veces el radio de cada cuerpo.

Otro pero que puede formularse al modelo de Safronov –y también a la concepción de Greenberg– es que trata a los planetesimales como si se estuvieran moviendo en trayectorias en línea recta, lo mismo que las moléculas en la teoría cinética de los gases. Concretamente, Safronov examina los planetesimales en una región del espacio en torno a un planetesimal dado. De aquí, obtiene la ventaja de que los planetesimales pueden describirse como si fueran partículas de un gas comprimido en un recipiente. Este método excluye un adecuado tratamiento de cualquier fenómeno que dependa de la distancia de los planetesimales al Sol.

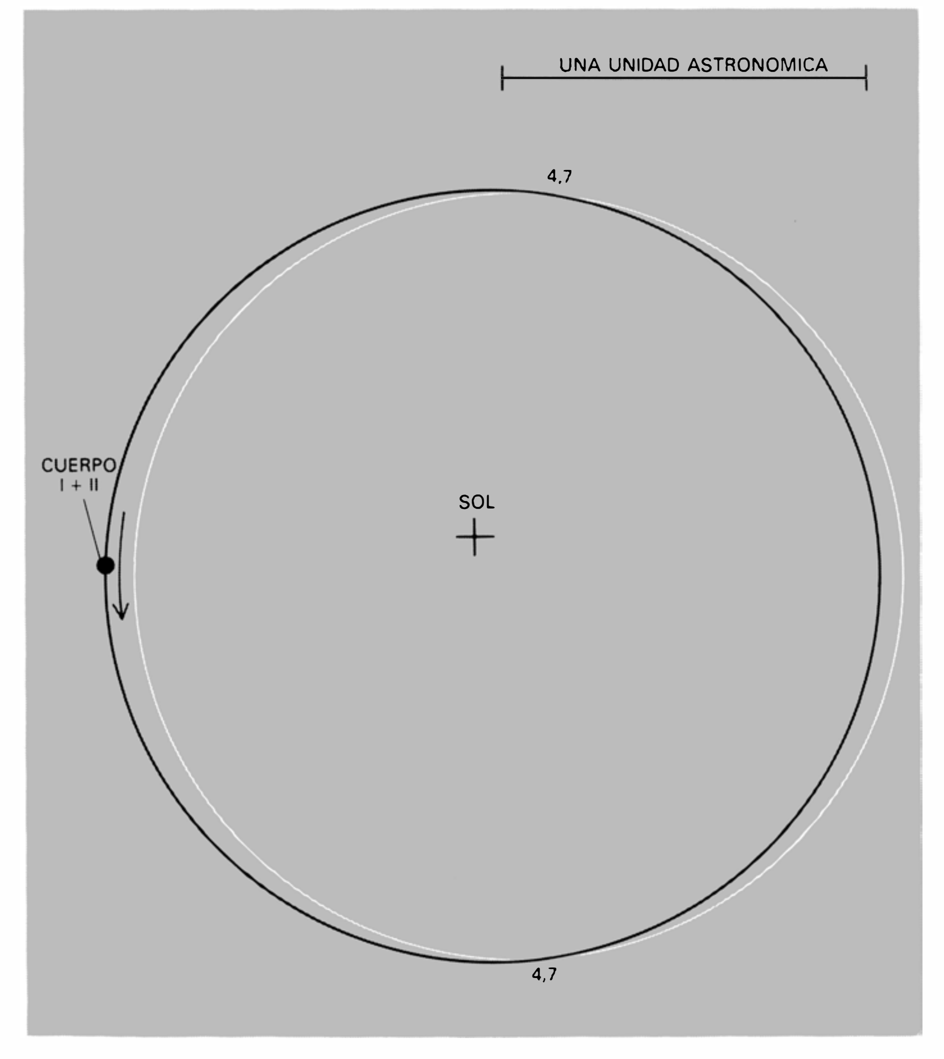
La duda fundamental acerca de las simplificaciones implícitas en el modelo de Safronov atañe a si la idea de un equilibrio entre la disminución de la velocidad causada por colisiones es válida en una teoría heliocéntrica. Cabe esperar que no sea así. Después de todo, la velocidad de un cuerpo en una órbita heliocéntrica depende de su distancia al Sol. Esto introduce una relación entre la distribución radial de los planetesimales y sus velocidades relativas, que no se da en la simple teoría cinética de los gases.

Yo he realizado estudios mediante ordenador que pueden ayudarnos a aclarar esta duda. El estudio comienza suponiendo un enjambre de 100 cuerpos de igual masa. Al principio, los planetesimales siguen órbitas de baja excentricidad. Los cambios orbitales resultantes de los acercamientos serían

pequeños, pero numerosos. Se puede calcular cada cambio, y determinar las nuevas órbitas. Las velocidades relativas y las excentricidades tienden al principio a incrementarse rápidamente, pero luego su tasa de incremento disminuye deprisa, y al final tienden a estabilizarse en valores proporcionales a la velocidad de escape. Este es el mismo efecto determinado por Safronov. Si se escoge un modelo inicial en que las órbitas tienen una elevada excentricidad, el proceso se opera en sentido opuesto: la velocidad relativa es grande al principio, y luego decrece. Al final se estabiliza en el mismo valor alcanzado en el enjambre de baja excentricidad.

El valor de la velocidad tras la estabilización basado en estos cálculos es alrededor de unos dos tercios el valor previsto por la teoría de Safronov. Esta coincidencia aproximada es todavía satisfactoria. La relación entre la velocidad relativa y la velocidad de escape queda confirmada, y en la medida en que hay una tendencia del sistema a regularse a sí mismo, la velocidad relativa se mantiene en un grado favorable a la coalescencia.

Los cálculos permiten también algo más que confirmar los anteriores resultados. Por ejemplo, podrían deducir que un enjambre, inicialmente confinado a un estrecho ámbito de distancias al Sol, se extiende hasta incluir un alto margen de distancias. Cada planetesimal toma parte en el incremento: se mueve en el creciente rango de radios orbitales, quedando unas veces cerca de un extremo, a veces cerca del otro y, las más, en la zona central. De hecho, el planetesimal toma un camino aleatorio en lo que se refiere a su distancia heliocéntrica. La oscilante órbita de cada planetesimal incrementa el número de oportunidades para llegar a la colisión con cuerpos distantes. Así, la excentricidad de las órbitas no es una condición necesariamente requerida para producir acercamientos. Aunque el efecto no elimina el riesgo de la formación de planetas menores demasiado pequeños y demasiado numerosos, reduce esta probabilidad.



COLISION DE DOS PLANETESIMALES. Los asocia en un solo cuerpo, que toma una órbita más circular. Se muestra la velocidad del cuerpo que se ha integrado, respecto del que no ha tomado parte en la colisión. Por término medio, las colisiones reducen las velocidades relativas de los planetesimales. Además, los planetesimales tienden a quedar aislados; en sus órbitas más circulares, tienen menos probabilidades de cruzarse con las órbitas de otros cuerpos, y, por consiguiente, de entrar en colisión.

Hasta aquí los cálculos se han ceñido al caso de cuerpos de igual masa. Estos se limitan también a determinar una situación de distribución estable de velocidades en un enjambre de planetesimales que pueden colisionar pero ya no engrosarse (aunque el ordenador calcula nuevas órbitas para los cuerpos que colisionan, no permite que se unan por coalescencia). Se está avanzando en la extensión de los cálculos al caso de planetesimales de diferentes masas.

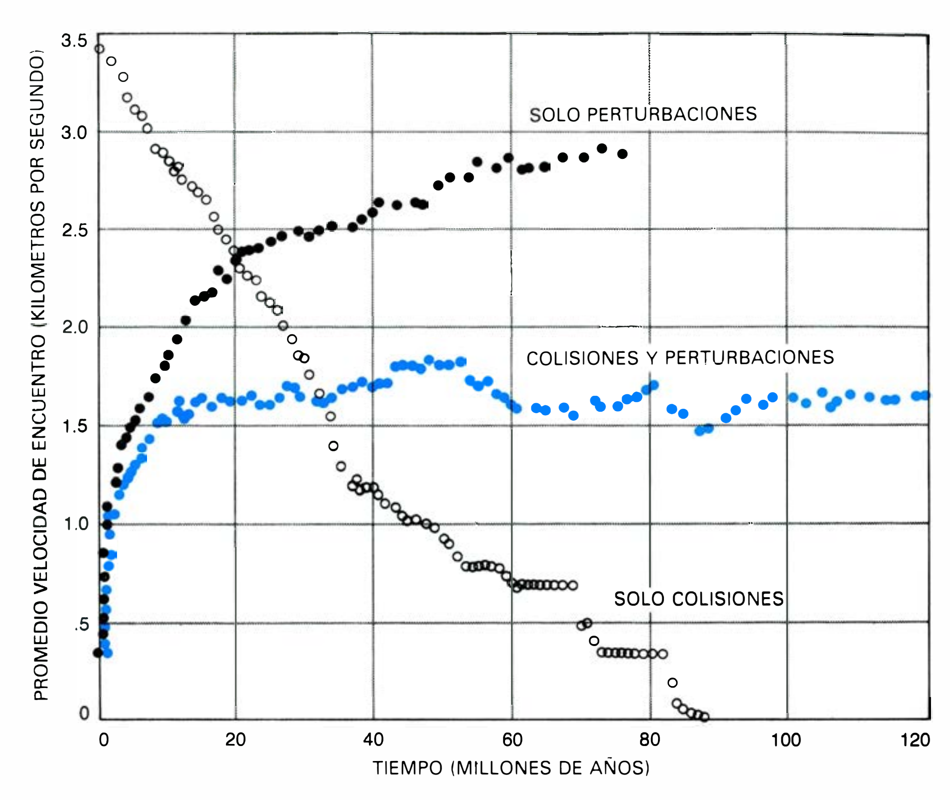
Un problema central es la necesidad de explicar las circunstancias que conducen a un enjambre de planetesimales a formar finalmente cuatro cuerpos –los planetas terrestres–, y no otro número cualquiera. De nuevo nos encontramos aquí con dos procesos contrapuestos. La ralentización colisional tiende a culminar el engrosamiento planetario con la formación de grandes cuerpos aislados moviéndose en órbitas casi circulares. Las perturbaciones gravitacionales operan el efecto opuesto: tienden a incrementar las excentricidades orbitales y a desplazar las órbitas radialmente, permitiendo la continuación de los acercamientos. Más tarde o más temprano, cuando todos los cuerpos grandes se hallen en órbitas que no se cortan, y todos los cuerpos pequeños que cruzan esas órbitas hayan sido capturados, la tendencia hacia unas órbitas casi circulares acabará por prevalecer. A no ser que la simulación demuestre la existencia de perturbaciones gravitacionales de largo alcance, no resta ya influencia alguna que pueda provocar la intersección de las órbitas de estos cuerpos grandes. La cuestión estriba en averiguar cuántos cuerpos grandes existen entonces.

Larry P. Cox, del Instituto de Tecnología de Massachusetts, ha realizado algunos cálculos sobre la formación simultánea de varios planetas. En su trabajo, se han supuesto las órbitas de 100 cuerpos de igual tamaño, distribuidos inicialmente al azar, en la región del sistema solar que ahora ocupan los planetas terrestres. Los cuerpos se perturban unos a otros, y al mismo tiempo colisionan y se engrosan. A la larga, sólo permanecen los cuerpos cuyas órbitas no se cortan. ¿Hasta qué punto los tamaños y distancias de los "planetas" finales recuerdan a los de los planetas terrestres?

Cox concluye que la respuesta depende del grado inicial de excentricidades. Cuando se parte del supuesto de

unas órbitas marcadamente elípticas, se obtiene un resultado final sorprendentemente parecido al sistema de los planetas terrestres. En una de las simulaciones, el número final de planetas era de seis. Uno de ellos era un planetesimal francamente cerca del Sol, que nunca podría colisionar con ningún otro cuerpo. Los otros planetas podrían llamarse, si queremos, Mercurio, Venus, Tierra, Luna y Marte.

Cuando se simulan bajas excentricidades en el estadio inicial –más de acuerdo con el modelo de Safronov– se forman por lo menos 10 planetas. Hasta cierto punto, la dependencia de la excentricidad resulta del hecho de que los cálculos de Cox se han limitado a un movimiento bidimensional. En este sistema solar "plano", las órbitas de los planetesimales se intersecan con frecuencia. Por tanto, la ralentización colisional (opuesta a la perturbación gravitatoria) es mayor que la que tendría lugar en un sistema de tres dimensiones. Mis propias simulaciones tridimensionales lo confirman. La ralentización colisional podría ser más importante en un enjambre cuyos miembros tuvieran diferentes tamaños iniciales. Un cálculo que parte de cuerpos de igual tamaño podría así conducir, de una manera u otra, a un aumento del número de colisiones respecto del de acercamientos. También se están haciendo cálculos en este sentido.



VELOCIDAD DE ACERCAMIENTO MEDIA en un enjambre de planetesimales. En un tiempo dado está determinada por el balance de dos efectos: los simples acercamientos tienden a incrementar las velocidades relativas, mientras que las colisiones tienden a reducirlas. Aquí se dan los resultados del cálculo de velocidades relativas de 100 cuerpos. A cada cuerpo se ha asignado una masa de 9×10^{24} gramos y un radio de 830 kilómetros. Se han escogido órbitas iniciales de baja excentricidad. Después de cada colisión o acercamiento, el ordenador recalcula las órbitas de los cuerpos que han intervenido. El gráfico muestra el promedio de la velocidad de encuentro bajo la influencia de los acercamientos y bajo la influencia de las colisiones. Cuando los efectos de ambos órdenes de interacción se han combinado, la velocidad se estabiliza en torno a un valor constante, de alrededor de 1,6 kilómetros por segundo.

La entera comprensión de la formación de planetas a partir de planetesimales requiere todavía investigar en otro terreno: la formación de los planetas en el seno de un medio gaseoso. Al fin y al cabo, la masa del sistema solar está formada, casi en su totalidad, por los gases altamente volátiles hidrógeno y helio, que son los principales componentes del Sol, Júpiter y Saturno. Estos gases impregnaban la primitiva nebulosa solar, y su papel está ya incluido en las hipótesis que describen cómo los planetesimales de un kilómetro de diámetro pueden haberse formado a partir de granos de polvo en choque unos con otros.

Puesto que el hidrógeno y el helio no son ahora los principales materiales constitutivos de los planetas terrestres, hay que suponer que deben haberse escapado de la región terrestre del sistema solar. Tal vez el primitivo viento solar era más bien un huracán solar. En todo caso, la presencia de los gases debe haber afectado a las primeras etapas de la acumulación de los planetas. Ciertas fuerzas adicionales han de haber actuado sobre los planetesimales mientras estos se movían en un medio gaseoso que ofrecía resistencia. La magnitud de estas fuerzas depende de la distribución del medio, el cual depende a su vez de los campos gravitatorios de los planetesimales o de los planetas, a medida que se van formando. Las simulaciones son, naturalmente, más difíciles que aquellas que prescinden de la existencia del gas. Sin embargo, este medio gaseoso ha sido incluido en modelos analíticos construidos por Chushiro Hayashi y sus colegas, en la Universidad de Kyoto. Su trabajo abarca no sólo los primeros estadios del engrosamiento planetesimal (los estadios ya descritos por Greenberg), sino también los más tardíos, en los cuales se formaron cuerpos del tamaño de la Luna, o incluso mayores.

La resistencia al avance opuesta por el gas varía de acuerdo con el área superficial del cuerpo que se mueve. El cambio de velocidad que resulta del efecto de una resistencia determinada es menor para una masa grande que para una pequeña, porque la grande tiene mayor inercia. El efecto de un medio gaseoso es así más importante para pequeños que para grandes cuerpos. El efecto sobre un cuerpo lo suficientemente pequeño es que su órbita se hace casi circular (un cuerpo que sigue una trayectoria circular sufre la mínima resistencia por parte del gas). Y no sólo esto, sino que el cuerpo tiende a describir espirales muy poco pronunciadas, acercándose paulatinamente al Sol. Durante un tiempo, este movimiento en espiral favorece el choque y engrosamiento de los planetesimales. Entonces, la mayor masa de los cuerpos que se han engrosado disminuye el efecto de la resistencia opuesta por el gas; de suerte que conforme los cuerpos ganan masa, su tendencia al movimiento espiral disminuye.

El resultado podría aparecérsenos como un enjambre de cuerpos moviéndose en órbitas aisladas y casi circulares. Esta expectativa, sin embargo, olvida las mutuas perturbaciones gravitatorias entre los planetesimales, las cuales se incrementan conforme los cuerpos crecen en tamaño. Las perturbaciones aumentan la velocidad de los planetesimales con respecto al medio gaseoso. De este modo se incrementa la resistencia al avance, pues esta fuerza crece en razón del cuadrado de la velocidad. Un planetesimal inmerso en un medio gaseoso puede continuar así su tendencia a describir espirales y a

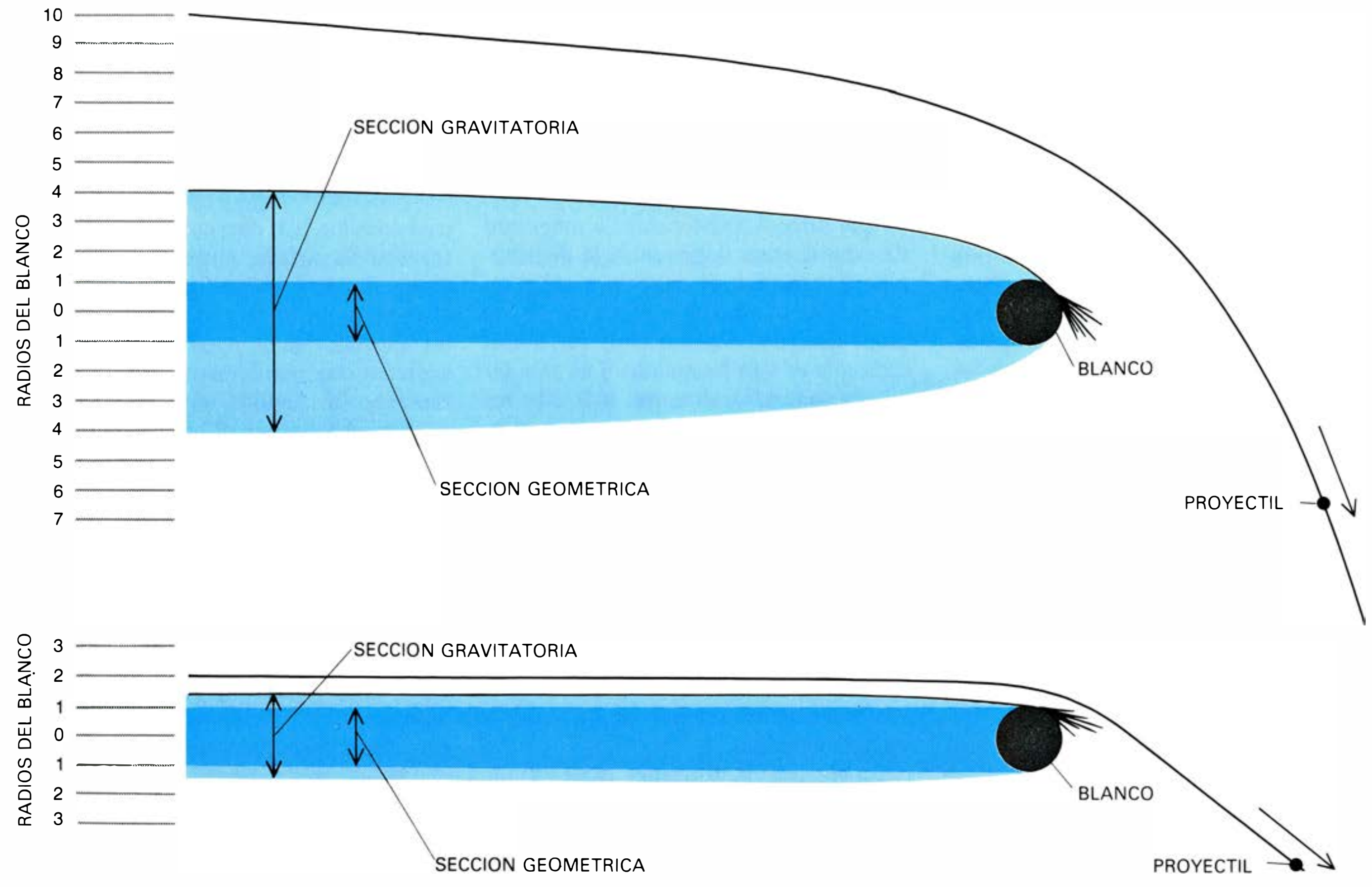
encontrarse con otros cuerpos, hasta alcanzar un tamaño de más de 1000 kilómetros.

En las simulaciones de Hayashi, cuerpos de un diámetro de 2000 kilómetros se habrían formado en 10.000 años. No está claro lo que sucede después. Es seguro que el medio gaseoso tiene una escasa influencia en la formación de cuerpos mayores. Estos deben encontrarse aislados unos de otros, en órbitas de baja excentricidad. Todo el resto de la masa del enjambre lo constituyen entonces cuerpos pequeños, que tienen órbitas excéntricas, con tendencia a la espiral. Si esos cuerpos llegan a colisionar y agregarse unos a otros, hasta resultar planetesimales con un diámetro entre los 100 y los 2000 kilómetros, las perturbaciones gravitatorias entre ellos podrían incrementar su velocidad respecto del medio gaseoso. La resistencia del gas, a su vez incrementada, haría más pronunciadas las espirales. Cabe imaginar que de este modo los planetesimales pueden emigrar radialmente hasta intersecar las órbitas de los cuerpos más grandes, es decir, los embriones de los futuros planetas. Sin embargo, si el movimiento migratorio es demasiado lento, ellos mismos pueden convertirse en embriones y adquirir órbitas aisladas casi circulares. El resultado podría ser un conjunto de planetas demasiado numerosos y demasiado pequeños.

Como vemos, es francamente difícil evaluar las probabilidades de estos acontecimientos. Hayashi ha previsto, sin embargo, la composición de la atmósfera que cada planeta terrestre recién formado pudo atraer hacia sí, detrayéndola de la nebulosa solar. Esta atmósfera incluye cierta proporción de gases inertes, tales como el argón 36. La concentración de tales gases puede medirse todavía en las rocas de la Tierra. Así, es posible averiguar si nuestro planeta se formó en el seno de un medio gaseoso, no sólo por teoría y simulación, sino por una forma directa de análisis.

Muchos geólogos quisieran poder comprobar por su cuenta que la Tierra primitiva fue realmente tal como la muestran las formulaciones físicas o los modelos de computación. Habrán, sin embargo, de tener paciencia. Con todo, ya en el actual estado de los trabajos teóricos, cabe demostrar que algunos de los posibles estados iniciales de la Tierra son más probables que otros. El primer caso es el referente a la temperatura inicial. Durante generaciones se ha discutido sobre si la Tierra se formó a una elevada temperatura, y por tanto estaba en gran parte fundida, o si fue siempre un cuerpo sólido y frío. Pues bien, las actuales teorías sobre formación planetaria prevén una alta temperatura inicial.

Consideremos primeramente las simulaciones de la acumulación de los planetas terrestres a partir de planetesimales en ausencia de gas. Durante los primeros estadios de acumulación, algunos de los planetesimales de un kilómetro llegaron a convertirse en cuerpos de 100, y luego de 1000 kilómetros de diámetro, a expensas de sus compañeros más pequeños. Tales cuerpos se formaron en la región del espacio que acabaría ocupando el sistema solar te-



ENFOQUE GRAVITATORIO, que es responsable de que un cuerpo grande pueda capturar a uno pequeño cuando podía pensarse que el pequeño no llevaba camino de colisionar con él. En el dibujo superior, la velocidad de dos pequeños cuerpos ("proyectiles") respecto de un cuerpo grande (el "blanco") es de 1,45 kilómetros por segundo. Uno de los cuerpos pequeños colisiona con el blanco; en ausencia del campo gravitatorio del blanco, hubiera pasado en su mínima distancia a 4,1 radios del blanco. El segundo cuerpo hubiera pasado a 10 radios del blanco. En vez de esto, su trayectoria respecto del blanco se ha hecho hiperbólica, y llega a acercarse a unos cinco radios del blanco. En el dibujo inferior, la velocidad de cada proyectil es de 5,8 kilómetros por segundo. El grado de enfoque gravitatorio se hace aquí menor. En las simulaciones de la formación del sistema solar interno, el enfoque gravitatorio puede dar origen a una distribución final en la que gran número de planetas relativamente pequeños han capturado a todos los planetesimales existentes y han adquirido órbitas que ya no se interfieren. En la ilustración, el "blanco" tiene una masa de 10^{27} gramos y mide un radio de 3970 kilómetros. Se supone que cada uno de los proyectiles es lo suficientemente pequeño como para que pueda despreciarse su perturbación gravitacional sobre el blanco.

rrestre. Ninguno de ellos tenía más probabilidades que cualquier otro de ser el embrión de un planeta. Safronov describe así la situación: "todos los planetesimales fueron creados iguales". Puede calcularse que sólo un pequeño número de planetesimales terminaron por adquirir una velocidad capaz de permitirles escapar del sistema solar. Por supuesto, sabemos que el sistema solar terrestre está hoy vacío, si exceptuamos la presencia de los planetas y sus satélites. Por tanto, casi todos los planetesimales llegaron a formar parte de los planetas, y el estadio final de la coalescencia de los planetas se caracteriza por impactos a alta velocidad entre cuerpos muy grandes.

Como Joseph Barrell, geólogo de la Universidad de Yale, observó en 1918, el impacto de tales cuerpos podría liberar gran parte de su energía en el seno del planeta naciente. Todos los fragmentos de la colisión podrían hacerse parte de ese planeta naciente; su energía cinética podría, por tanto, quedar capturada. Además, el calor generado por la colisión emplearía millones de años en atravesar 1000 kilómetros de tierra (del mismo modo que unos metros de tierra pueden preservar un bloque de hielo durante un verano). En efecto, el calor podría quedar enterrado junto con los escombros de la colisión. Curiosamente, la única energía que podría escapar inmediatamente de la colisión sería la relativamente pequeña cantidad de calor que los fragmentos arrojados fuera de la zona de choque pudieron irradiar al espacio antes de caer. Los cálculos sobre el efecto térmico de los impactos indican que los cuerpos así acumulados, si son algo mayores que la Luna, estarían fundidos en alguna extensión, y cuerpos del tamaño de la Tierra estarían parcialmente fundidos a todo su través.

La formación de la Tierra en el seno de un medio gaseoso también supone colisiones entre grandes planetesimales. Hayashi ha demostrado, además, que en un medio gaseoso cada planeta en formación atrae una masiva atmósfera inicial. La compresión de la atmósfera, por gravedad, en una Tierra naciente, pudo por sí misma conducir a una temperatura más alta que el punto de fusión de rocas, mucho antes de que el planeta estuviera formado.

La única alternativa a la formación de la Tierra por acumulación de planetesimales es el desarrollo de una profunda inestabilidad gravitatoria en la nebulosa solar, de modo que la Tierra cayera sobre sí misma por efecto de su propio peso. A. G. W. Cameron, del Centro de Astrofísica del Observatorio de Harvard y del Observatorio Astrofísico Smithsoniano, ha mostrado que este proceso también supone una alta temperatura inicial.

Durante varias décadas, una serie de investigadores, entre los que se incluyen Harold C. Urey, William W. Rubey y A. P. Vinogradov, han presentado argumentos en que se sostiene una tesis contraria. Un más reciente defensor de esta postura es A. E. Ringwood, de la Universidad Nacional Australiana. Basa su argumentación en que podemos considerar la Tierra como una enorme refinería química. Si una parte de su interior alcanza una temperatura dada, ciertas sustancias se concentran en una roca fundida y otras quedan sueltas como residuos sólidos. Separaciones químicas similares dan lugar a cristales de minerales puros cuando la roca se vuelve a solidificar de nuevo. Las separaciones químicas asociadas a la fusión y resolidificación son la razón principal de que la corteza continental sea distinta de la del fondo de los océanos. Si la Tierra estuvo más o menos fundida en un principio, se puede argüir que las separaciones que produjeron los continentes se habrían completado hace mucho tiempo. La corteza continental debería ser más extensa de lo que es, y no se hubiera podido seguir formando nueva corteza.

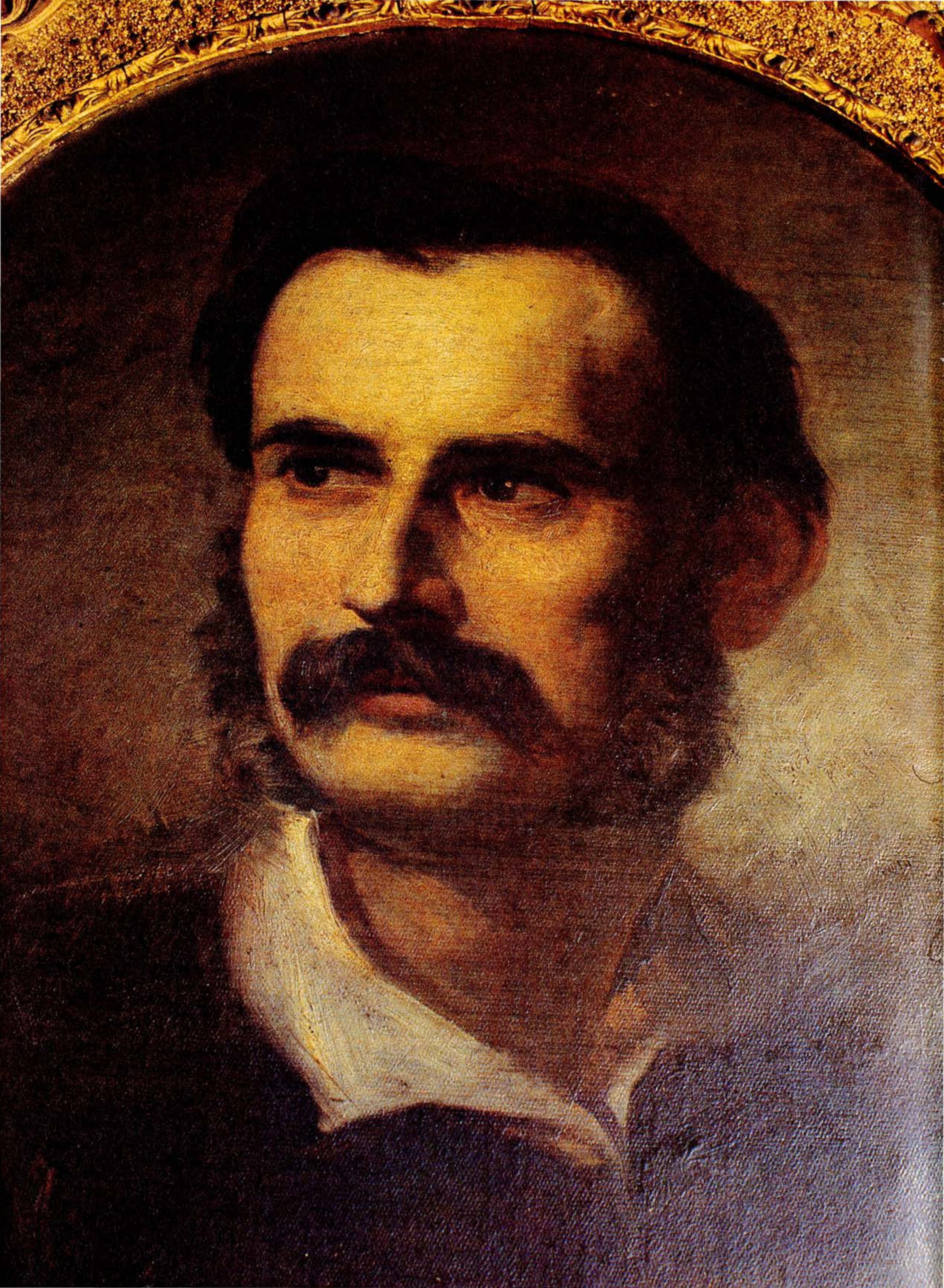
¿Por qué, entonces, las diferenciaciones químicas de la corteza continúan todavía? La cuestión espera una correcta respuesta. De momento, sin embargo, esta respuesta ha de ser vaga. El calor generado en la Tierra todavía hoy por la desintegración de núcleos radiactivos de vida larga se transporta hasta la superficie primeramente por convección en el manto terrestre. Parece que sólo una pequeña fracción de la roca convectiva está fundida. Se sabe que un pequeño grado de fusión favorece la separación química. Los grandes iones, por ejemplo, se escapan con facilidad de un conglomerado de cristal parcialmente fundido. Por contra, la fusión de una roca entera deja toda la masa fundida.

Se ha calculado que el impacto de planetesimales puede haber calentado la Tierra por coalescencia, en una proporción mil veces mayor que hoy la radiactividad. En tanto la Tierra se formaba, la convección debió haber sido mucho más rápida. Una gran parte de las rocas habría estado fundida, y es verosímil que el transporte de calor a través de la roca fundida fuese tan importante como el transporte por convección en estado sólido. Imaginemos que la velocidad de convección era rápida, que el grado de separación química era pequeño y que la densidad de las rocas en solidificación en la superficie no era mucho menor que la densidad del manto subyacente. En esas condiciones, cabe deducir que la recién formada corteza de la Tierra simplemente quedase incrustada en la zona interior. Tal vez, en otras palabras, la tectónica de la Tierra debió evolucionar despacio, y la cuantía de la fusión habría sido pequeña si las rocas de baja densidad que caracterizan los continentes fueron apartadas del resto y han quedado aisladas y estables.

Otro tema de importancia geológica es el poder precisar a partir de la teoría cuánto tiempo tardó en formarse la Tierra. La formación de la Tierra por efecto de la inestabilidad gravitatoria en la nebulosa solar requiere alrededor de 100.000 años. La formación de la Tierra a partir de planetesimales en ausencia de un medio gaseoso exige 100 millones de años. Sin embargo, en sólo 20 millones de años estaba ya ensamblada la mitad de su masa. Aunque la formación de la Tierra a partir de planetesimales en el seno de un medio gaseoso es menos determinable, parece poco probable que la presencia de este medio modificara mucho ese intervalo temporal.

En contraste con todas estas previsiones teóricas, las concentraciones relativas de varios isótopos del plomo en los minerales terrestres sugieren (de acuerdo con la interpretación más simple de los datos, y suponiendo que la Tierra estaba caliente cuando se formó) que el proceso fue notoriamente más lento. Incluso se pretende que después de 100 millones de años la Tierra estaba integrada en sólo su mitad. Todos los modelos teóricos de la formación de la Tierra arguyen, sin embargo, contra esta simplista interpretación de los isótopos. Tal vez un día conoceremos una más completa, aunque no menos plausible, interpretación de la historia química de la Tierra. Una ciencia sana se caracteriza por el continuo intercambio entre la teoría y las observaciones.

Hasta ahora los geólogos han visto probablemente las teorías sobre la formación de la Tierra basculando en un campo que va desde la especulación descarada hasta un inocente entretenimiento. Sin embargo, hoy se están convirtiendo en una seria empresa. La investigación geológica requiere, tanto como las misiones espaciales, suposiciones teóricas acerca de cómo se ha formado la Tierra. No podemos esperar que el proceso de nuestros conocimientos se opere demasiado deprisa.



Los "Ictíneos" de Narcís Monturiol

Por las innovaciones introducidas y por la experiencia acumulada en las numerosas pruebas realizadas, representaron la resolución casi definitiva de los problemas planteados hasta entonces en la navegación submarina

Santiago Riera i Tuebols

Si pasamos por alto las campanas de inmersión, ya conocidas en la antigüedad –Aristóteles nos habla de ellas y Alejandro Magno las utilizó en el asedio de Tiro el año 332 a. de J.C.–, el primer intento serio de navegación submarina corresponde a Cornelius Drebbel (1572-1633), quien, en el año 1620, recorre las dos millas existentes entre Westminster y Greenwich, por el Támesis, con una tripulación de quince hombres, de los cuales doce, aplicados a los remos, constituían la fuerza motriz que hacía avanzar la nave a una profundidad que oscilaba entre los 3,5 y 4,5 metros. La inmersión se efectuaba admitiendo agua como lastre en depósitos situados convenientemente; y la emersión, soltando pesos.

Veinte años después, Marin Mersenne proyecta un nuevo tipo de submarino. Más tarde, John Wilkins (1614-1672), obispo de Chester, vuelve a ocuparse del mismo tema en su *Mathematical Magick* (1648). En 1683, Giovanni A. Borelli ideó el procedimiento de provocar la inmersión admitiendo agua en unos conductos o mangas, de manera que, al ser comprimidos, se expulsara aquella de nuevo, originando así la emersión de la nave. También Denis Papin construyó su barco sumergible, interesándose por el problema de la renovación de aire, que solucionó con un tubo cuyo extremo libre se mantenía fuera del agua mediante un flotador, al mismo tiempo que pensaba en la nave como máquina de guerra. Precisamente por intuir que un buque de tal naturaleza podía convertirse en un arma atroz, Leonardo de Vinci destruyó los planos de su artilugio submarino.

Noventa años después de la prueba efectuada por Borelli, J. Day permanece 24 horas a diez metros de profundidad. Pero Day será la primera víctima de esta historia cuando al año siguiente intente descender a más de cien metros y no torne a la superficie. En 1776, el norteamericano David Bushnell, en plena guerra de la independencia de los Estados Unidos, idea un extraño submarino de forma ovoidal que, mediante palancas y pedales que accionan una hélice, puede avanzar con lentitud y a escasa profundidad: se trata del *Turtle*. En la primera acción de guerra propiamente dicha de la historia submarina moderna, se acercó al buque inglés *Eagle* para fijar un explosivo en su casco, acción que fracasó por mor de las planchas protectoras de cobre que recubrían el cuerpo del navío.

El siguiente intento de importancia corresponde a Robert Fulton. Napoleón prestó oídos a sus explicaciones y convirtió en realidad su *Nautilus*. Navegaba a vela en superficie y, mediante una hélice accionada a mano, en inmersión. De todas formas, Fulton no supo –o no pudo– despertar el entusiasmo necesario y, medio olvidado, olvidó él también la nave submarina para dedicarse al estudio y proyecto de un barco de vapor. P. Antoine Payerne se interesó por el problema de la purificación del aire en locales herméticamente cerrados, y llegó a construir una nave submarina conocida con el nombre de buque de Payerne. La relevancia de este médico francés nacido en 1806 residen en el influjo que ejerció en Monturiol, no en sus logros técnicos.

Finalmente, el alemán Wilhem Bauer construyó en los astilleros de Kiel un pequeño submarino de 7,9 metros de longitud; pero se hundió al efectuar la primera prueba, salvándose sus tripulantes, entre los que se encontraba el propio Bauer: al darse cuenta del accidente, abrieron la escotilla y salieron catapultados por la presión del aire que previamente habían encerrado en el barco. Volvió a la carga con otro sumergible, el *Diablo del mar*, construido en Kronstadt el año 1856; iba accionado por pedales y podía permanecer varias horas bajo la superficie del mar. Tampoco le acompañó el éxito: el *Diablo del mar* se hundió también al año siguiente. En 1863, los sudistas estadounidenses emplearon campanas sumergibles para combatir a la flota de la Unión. Pero Monturiol, en Barcelona, había probado ya su *Ictíneo I*.

Narcís Monturiol y Estarriol nació en Figueres (Gerona) el 28 de septiembre de 1819. Vivió una agitada juventud de activa participación en la política del país, hecho que explica los sucesivos confinamientos –forzados o voluntarios– fuera de su tierra natal. Abogado –profesión que nunca ejerció–, cajista –oficio que más tarde le valió la subsistencia– y pintor eventual, dedicó gran parte de la primera mitad de su vida al periodismo, fundando y colaborando en diversas publicaciones.

Fue uno de los introductores y defensores, en España, de la doctrina del socialista utópico Etienne Cabet, con quien mantuvo una relación directa. Consiguió aglutinar a su alrededor, en Cataluña, un grupo de simpatizantes "cabetianos" que, aunque modestamente, contribuyeron en dinero y hombres a la famosa aventura de la fundación de la idílica Icaria. Fallido el intento, y de nuevo activamente integrado en la política interna, se vio forzado a refugiarse en el pueblecito costero de Cadaqués, donde, contemplando el enorme esfuerzo de los hombres que se

NARCIS MONTURIOL Y ESTARRIOL (1819-1885), inventor, diseñador de los dos *Ictineos* y político, en un retrato del pintor catalán Ramón Martí y Alsina, pintado en la década de 1850, probablemente durante su exilio voluntario en Cadaqués (Gerona). Se halla en el Museo Marítimo de Barcelona.

dedicaban a la pesca del coral, empezó a pensar en un artefacto que suavizara su esfuerzo y a la vez aumentara la productividad de la operación.

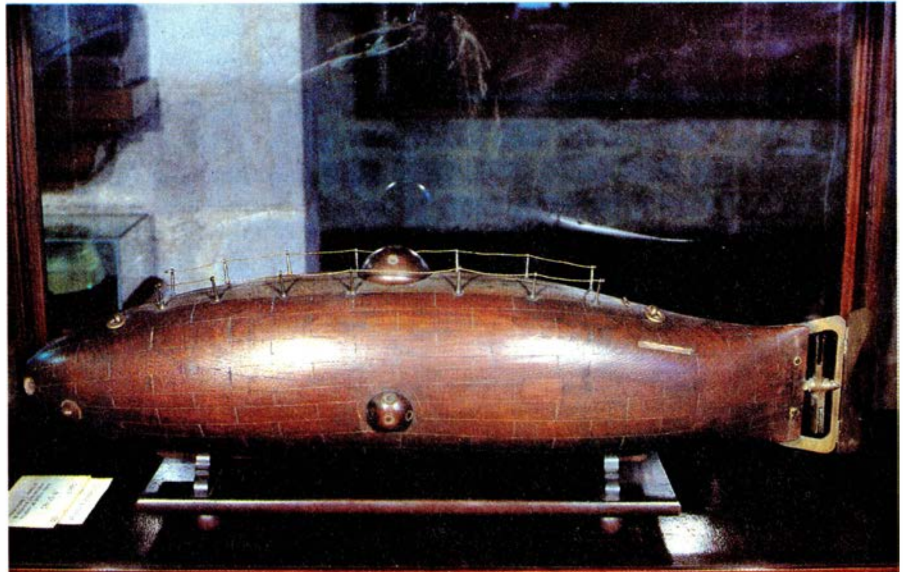
De vuelta a Barcelona, trabaja con ahínco en esa idea y se apresta a construir un navío que "navegue por debajo del agua". Convence a un grupo de conocidos, amigos e interesados en el proyecto, entre los que se encuentra Martí Carlé, para formalizar la escritura de un capital de 2000 duros, con los que empezará a levantar la nave, a la vez que se estipula en el mismo documento el sueldo de Monturiol: treinta reales diarios más la tercera parte de todos los beneficios que puedan obtenerse en la utilización del submarino.

La recaptación del dinero necesario para la realización de un proyecto tan ambicioso continúa, y se suma a esta aportación inicial, mientras los trabajos iniciados van dando forma al "barco-pez" en los talleres de "El Nuevo Vulcano", en el barrio de la Barceloneta. En octubre de 1858 la cantidad reunida sube ya a la nada despreciable cifra de 73.900 reales. El 15 de abril de 1859 se recibe la autorización para efectuar las primeras pruebas.

Bautiza el barco con el nombre *Ictíneo* –del griego, *Ichtys*, pez, y *naũs*, nave–. Monturiol deseaba que su comportamiento fuese idéntico al del pez que respira, vira, se desplaza, sube o baja en el mismo medio. Este primer buque, *Ictíneo I*, fue construido con una idea pura y simplemente demostrativa: la de que era factible navegar sumergido en el agua sin peligro, o con el menor posible, y en ausencia de atmósfera terrestre. Y ello, durante largos períodos de tiempo.

Medía la nave siete metros de eslora, tres y medio de quilla a cúpula y dos y medio de manga. Poseía dos cascos; uno: la cámara resistente, cilíndrica y de sección elíptica; el segundo, en "forma de pez", según propias palabras, en la creencia de que la forma natural del modelo escogido en la naturaleza era la más apta para la navegación. La cámara de presión, o resistente, de siete metros cúbicos de capacidad, podía albergar cinco hombres: cuatro de ellos encargados de la tracción de la nave, aplicando su esfuerzo muscular para conseguir el giro de la hélice motriz, y el timonel, responsable de la inmersión, viradas y equilibrio de la nave. Dicha cámara se construyó con cuadernas de madera de olivo, reforzadas mediante un forro de cintas de cobre, recubriéndose el conjunto con un forro impermeable, también de cobre. Se lograban así las condiciones requeridas de resistencia, rigidez e impermeabilidad.

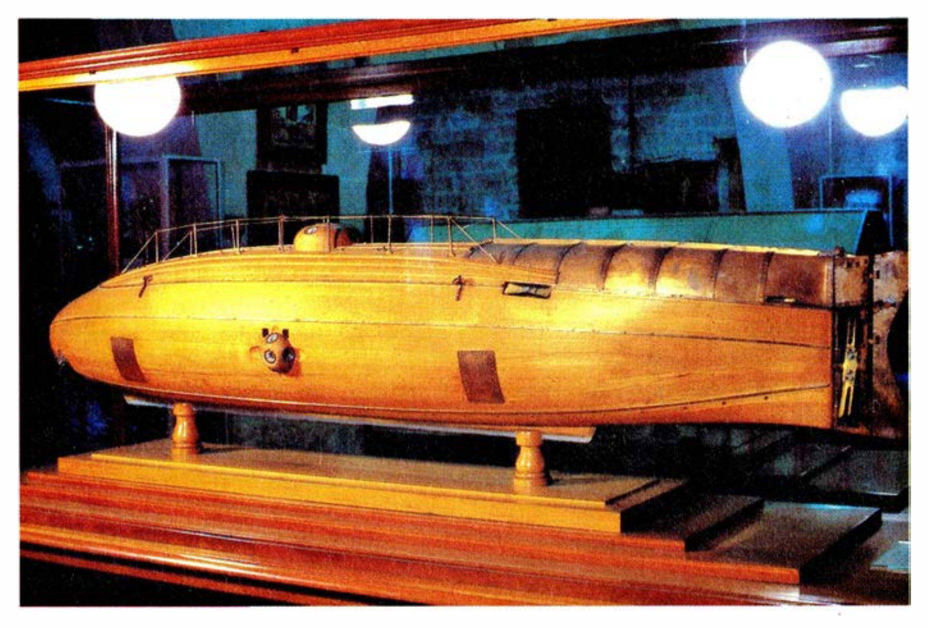
El hecho de haber escogido la madera como material básico obedecía a dos razones: era barata y la industria metalúrgica se encontraba atrasada en nuestro país. No obstante, es obligado puntualizar que el *Ictíneo I* se había calculado y proyectado para navegar a una profundidad de 500 metros, diez veces mayor que la máxima que se esperaba alcanzar, cifrada alrededor de los 50 metros. Tenía, pues, un margen de seguridad notable.



MAQUETA DEL ICTINEO I. Puede observarse la forma exterior de pez que tiene la nave y que su diseñador le confirió voluntariamente. Como dichos vertebrados, posee el aparato propulsor en la cola y está dotado de vejigas de flotación para iniciar la inmersión. En cuanto a su resistencia, el casco podía soportar una presión de ocho atmósferas que viene a ser la presión existente a 70 metros de profundidad.

Importantes eran, en el funcionamiento del submarino, los llamados mecanismos natatorios; en particular, las vejigas de flotación, compartimentos de cobre situados entre ambos cascos y repartidos a babor y estribor, inicialmente llenos de aire a baja presión. Cuando se deseaba sumergir la nave, bastaba abrir los grifos de admisión de agua y de escape de aire, procurando hacer la operación de modo ordenado para evitar que el buque escorase. Con los depósitos llenos, el *Ictíneo* tenía la densidad del agua. Quedaba así entre dos aguas, sumergido hasta la escotilla. Para subir o bajar, Monturiol idearía dos procedimientos: el empleo de vejigas natatorias o de presión (caso del segundo *Ictíneo*) y la hélice horizontal (en este primer *Ictíneo*), solución más barata aunque no del agrado del inventor.

La propulsión se obtenía por medio de la fuerza muscular de los tripulantes,



MAQUETA DEL ICTINEO II. Su forma sigue recordando a la de un pez, pero con notables variaciones respecto a su hermano menor. (Como la anterior, se halla en el Museo Marítimo de Barcelona.)

que accionaban y hacían girar, a través de una transmisión ordinaria, una hélice Mangin, de palas planas. Por este procedimiento, se alcanzó, en inmersión, una velocidad que podríamos establecer entre 0,5 y 1 metro por segundo, en el bien entendido de que esta última constituiría un límite superior. Me inclino a creer que el valor promedio oscilaba alrededor de un nudo, es decir, 0,5 metros por segundo; ignoro la velocidad alcanzada por el *Ictíneo I* en la navegación por la superficie, aunque no creo que superara la velocidad de 1 metro por segundo. Valores, estos, que pueden considerarse aceptables si tenemos en cuenta, además, las notables fuerzas de rozamiento existentes tanto por las proporciones de la nave como por su forma.

Otro problema que Monturiol se planteó fue el de la respiración en las largas inmersiones. Con respecto a este punto decidió llevar oxígeno embotellado a presión. Ahora bien, a mediados del siglo XIX, semejante providencia resultaba difícil y cara; de ahí que las inmersiones se acabaran realizando, de un modo un tanto temerario, solamente aprovechando al máximo el aire inicial del recintó herméticamente cerrado, sin ningún tipo de reposición ni purificación. "Así es que llevábamos el oxígeno almacenado en depósitos; y como era enojoso producirlo, comprimirlo y llevarlo al *Ictíneo*, muchas veces preferíamos abandonarnos al azar de las pruebas, la duración de las cuales procurábamos disminuir en lo posible, sin otro aire que el contenido inicialmente en nuestra cámara y sin otros instrumentos de análisis que los mismos efectos producidos en nuestra organización por el aire impuro que circulaba disuelto en nuestra propia sangre." Son palabras de Monturiol, quien prosigue en la relación de una de estas pruebas: "en una cámara en que había 1,166 metros cúbicos de aire por persona [éstas eran seis y la cubicación de la atmósfera del *Ictíneo* apenas llegaba a siete metros cúbicos], a la media hora, la presión barométrica había subido 12 milímetros; el termómetro subía algo hasta alcanzar, a las dos horas, 5 grados [se refiere al incremento de temperatura]; en este momento empezaban las molestias, la frecuencia de la respiración, la cianosis y la opresión, que obligaban a dar por terminado el ensayo al cabo de media hora. Mucho tiempo antes, a los 75 minutos de empezar la prueba, si hay una luz a bordo se extingue, de manera que con esto se tiene una guía segura que suprime otros aparatos que serían delicados y tal vez expuestos a errores. La atmósfera se hizo irrespirable al alcanzar un 3,85 % de ácido carbónico y al perder 5,02 % de oxígeno". Estas dificultades obligaron al replanteamiento y solución del problema, de una manera definitiva, en el segundo *Ictíneo,* como veremos.

CARACTERISTICAS	ICTINEO I	ICTINEO II
Eslora	7,0 metros	17,0 metros
Manga	2,5 metros	3,0 metros
Altura (desde la quilla hasta el extremo superior de la cúpula)	3,5 metros	3,5 metros
Sección del casco de presión	elíptica	circular
Longitud del caso de presión	~ 5,0 metros	14,0 metros
Capacidad interior del caso de presión	7 metros cúbicos	29 metros cúbicos
Desplazamiento volumétrico en la superficie	~ 10 metros cúbicos	64 metros cúbicos
Capacidad de los tanques de inmersión (vejigas de flotación)	—	8 metros cúbicos
Desplazamiento en la inmersión	—	72 metros cúbicos
Velocidad en la superficie	1 m/seg	1 – 2 m/seg
Propulsión en la superficie	4 hombres	16 hombres/6 CV
Velocidad en la inmersión	0,5 – 1 m/seg	0,5 – 1 m/seg
Propulsión en la inmersión	4 hombres	16 hombres/2 CV
Profundidad de inmersión proyectada	40 metros	50 metros
Profundidad de inmersión alcanzada	20 metros	30 metros
Número de inmersiones	54	13
Duración máxima de la inmersión	2 1/2 horas	7 1/2 horas
Número de tripulantes (capacidad máxima)	5	20

COMPARACION DE LAS PRINCIPALES CARACTERISTICAS de ambos *Ictíneos*. El móvil que estimuló a Monturiol a construir esos artefactos fue suavizar el esfuerzo que desarrollaban los pescadores de coral y les permitiera elevar su rendimiento. El primer *Ictíneo* se botó en 1859 y el segundo en 1864.

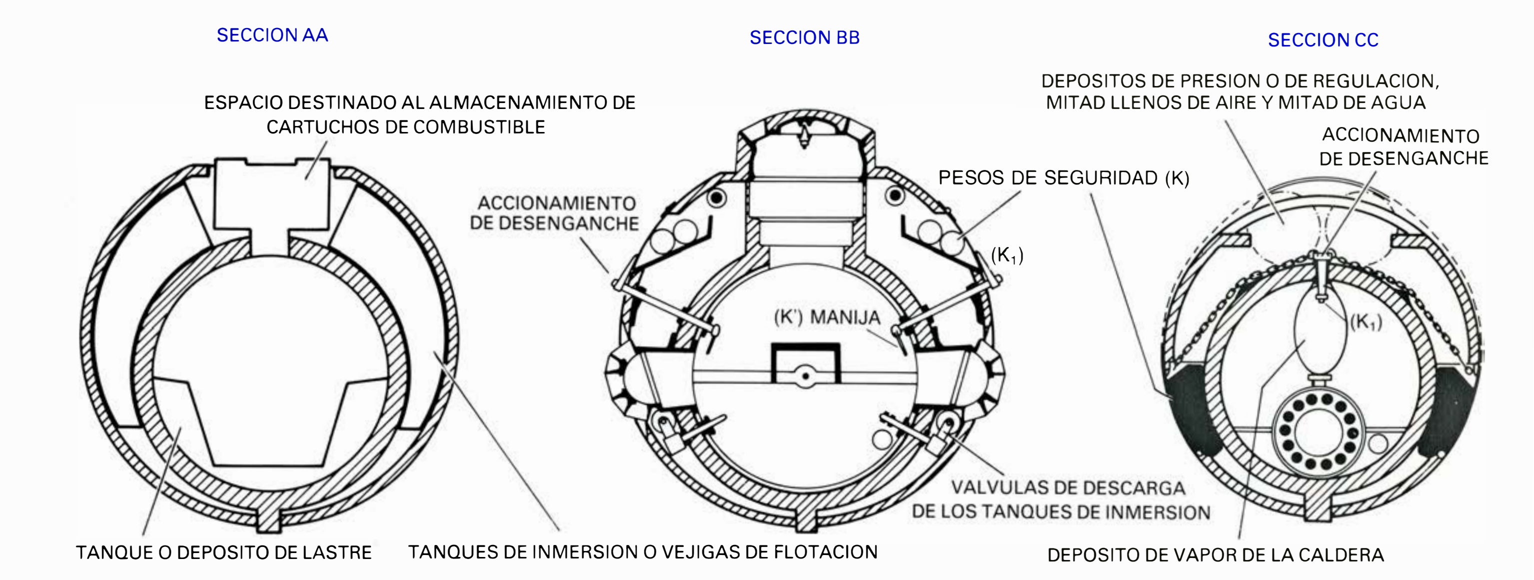
Acompañaron a Narcís Monturiol,en la primera prueba, Josep Missé y Castells, maestro de ribera y constructor del casco del primer *Ictíneo*, y Josep Oliu y Juan, tripulante. A medida que las pruebas se fueron sucediendo se sumaron otros técnicos: nueve en total, aparte de los ya mencionados, entre los que cabe citar explícitamente a los hermanos Ros y Bondia.

La botadura tuvo lugar el 28 de junio de 1859. Con poco éxito, pues el navío sufrió daños que afectaron a las vejigas de flotación, los forros y algunos cristales. "La avería fue de consideración y se me habían acabado los recursos pecuniarios. Pero, después de un examen detenido me convencí de que no había ningún inconveniente en efectuar las inmersiones mientras no superasen la profundidad de veinte metros." Se efectuaron las pruebas, con mejor resultado esta vez: "El Ictíneo sube y baja, camina y gira en la superficie, entre dos aguas y en el fondo del mar. El hombre vive tan bien dentro del Ictíneo como en plena atmósfera. La navegación submarina es una realidad". El éxito de estas primeras inmersiones le animaron a realizar una prueba pública, ante autoridades y periodistas, el 23 de septiembre del mismo año.

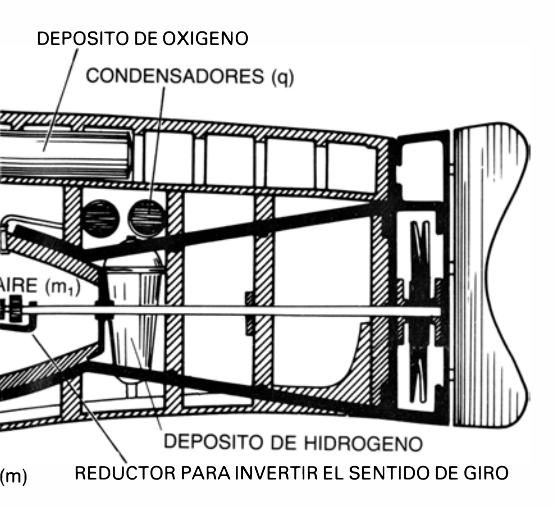
El 7 de mayo de 1861, y tras múltiples avatares (falta de dinero, oferta del invento al gobierno, viajes a Madrid, correspondencia con los organismos públicos, etcétera), se realizó una prueba oficial en Alicante (Alacant), con la asistencia de los ministros de Marina y Fomento, Zavala y Cervera, diputados, senadores y miembros de la Academia de Ciencias de Madrid. A pesar del mar de fondo, olas de cuatro metros y fuerte viento de levante, el *Ictíneo* demostró que la navegación submarina no era ningún sueño de loco. Barcelona y su ciudad natal, Figueres, reciben apoteósicamente a Narcís Monturiol.

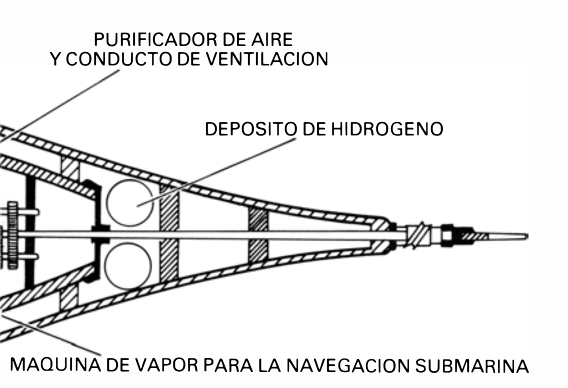
El *Ictíneo I,* proyecto en el cual se gastaron 20.000 duros, no había sido más que el principio. Monturiol pensa-

CONDUCTO DE AIREACION O RESPIRADERO (l)
A
VENTANAS
B
SALIDA DE AIRE DE LOS TANQUES DE INMERSION
C
D
ACCIONAMIENTO TIMON
CHIMENEA (n_2)
ALMACENAMIENTO DE CARTUCHOS DE COMBUSTIBLE
DEPOSITOS DE PRESION O DE REGULACION (g_1)
FAROS
BAROMETRO
(n_1)
PURIFICADOR DE ...
CONDUCTO DE VENTILACION (m_1)
TANQUE O DEPOSITO DE LASTRE (g)
BOMBA DE TRES CUERPOS (g_2)
CALDERA (n)
VENTILADOR
MANOMETRO DEL TANQUE DE REGULACION
MANOMETRO REGULADOR DE LA PROFUNDIDAD
BOMBA DE AIRE
BOMBA DE ALIMENTACION DE LA CALDERA
DEPOSITO DE VAPOR DE LA CALDERA (20 ATMOSFERAS) (n_1)
BOMBA DE AIRE Y CONDUCCION
PESO EQUILIBRADOR EN SENTIDO LONGITUDINAL
RECIPIENTE DE RETENCION DEL CO_2
TANQUES DE INMERSION O VEJIGAS DE FLOTACION
VALVULAS DE DESCARGA DE LOS TANQUES DE INMERSION
MAQUINA MOTRIZ PRINCIPAL: MAQUINA DE VAPOR ALTERNATIVA DE DOS CILINDROS (p)

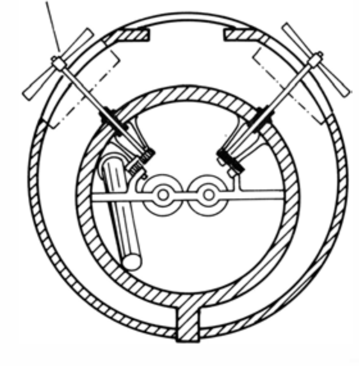


PLANO DEL ICTINEO II con la descripción de sus más importantes elementos. En él se distinguen el alzado, la planta y las secciones perpendiculares al eje longitudinal AA, BB, CC y DD en la parte inferior, situadas en el alzado mediante líneas de puntos. Observando con atención podemos visualizar y situar el sistema motriz: caldera, depósito de vapor y la máquina motriz propiamente dicha que acciona la hélice impulsora; el sistema de regeneración del aire con el ventilador, el purificador y el conducto de ventilación; también el sistema de inmersión: vejigas de flotación, bomba, depósito de lastre y depósitos de presión. Asimismo pueden situarse las dos máquinas que llevaba la nave: la motriz principal para la navegación en superficie y otra,









de menor potencia, para la inmersión. Obsérvese la multitud de elementos que obligaron, dada la precariedad del espacio, a un estudio racional de su colocación. [Del "Ensayo sobre al arte de navegar por debajo el agua" de Narcís Monturiol.]

ba ya en una segunda nave mucho más capaz. Empero, para concretar y realizar esta ilusión, hacían falta recursos monetarios que no poseía. Del gobierno sólo obtiene palabras y promesas, a pesar de que presenta su proyecto como barco de guerra, al que es posible adaptar cañones y convertirlos en lanzadores de torpedos, sin abdicar de los fines propuestos inicialmente: "si la observación casi exclusiva de la parte sólida de nuestro globo es la fuente de las ciencias físicas y químicas, la investigación de la profundidad de las aguas nos proporcionará nuevos datos que deberemos a estas mismas leyes de la naturaleza, la acción de las cuales será diferente pues se cumplirá en circunstancias también diferentes"... "Si, aparte de estas consideraciones generales que expresan mis presentimientos sobre la utilidad científica de la navegación submarina, nos fijamos en las especialidades ya citadas, en el estudio de las corrientes y volcanes de aquella vastísima región y en la distribución de la vida y de la luz, tendremos que convenir que la navegación subacuática incluye un interés muy elevado de actualidad, que trasciende del simple vehículo submarino." Decepcionado, opta, por fin, por la suscripción popular.

El *Ictíneo II* empezó a construirse el año 1862. Se constituyó la sociedad comanditaria "Navegación submarina" (razón social: Monturiol, Font, Altadill, C^{a}), para financiar la empresa, con un capital inicial de 58.900 duros, a los que había que sumar otros 20.000 duros procedentes de la suscripción popular que se había extendido a Cuba. El 2 de octubre de 1864 se botó el casco de la nueva nave, construido bajo la dirección de Joan Monjo y Pons, arquitecto naval.

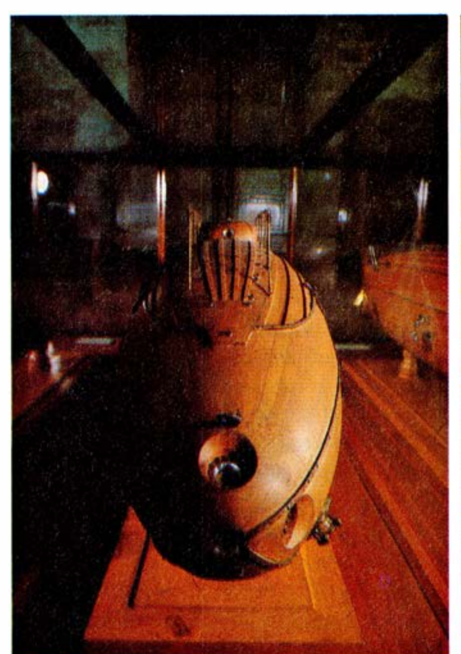
Medía 17 metros de eslora, tres y medio desde la quilla hasta la parte más alta de la cúpula y tres de manga. El casco interior o de presión tenía, en sentido longitudinal, forma elipsoidal y su volumen era de 29 metros cúbicos, notablemente mayor que el de la primera nave, estando formado por cuadernas de madera de olivo de 10 centímetros de espesor revestidas externamente con cintas adicionales de madera de roble, de 6 centímetros de espesor; todo el conjunto iba forrado con chapa de cobre de 2 milímetros. El casco estaba rematado en su parte delantera por una pieza de fundición de bronce, con ventanas, y por su parte posterior con otra del mismo material con el prensaestopas para el árbol de la hélice. Las chapas de cobre del casco de presión se soldaron con esmero y perfección raras en aquella época, convirtiendo así la nave en un cuerpo completamente estanco. Los dos cascos –el de presión y el exterior– se construyeron sobre una quilla de madera revestida de bronce.

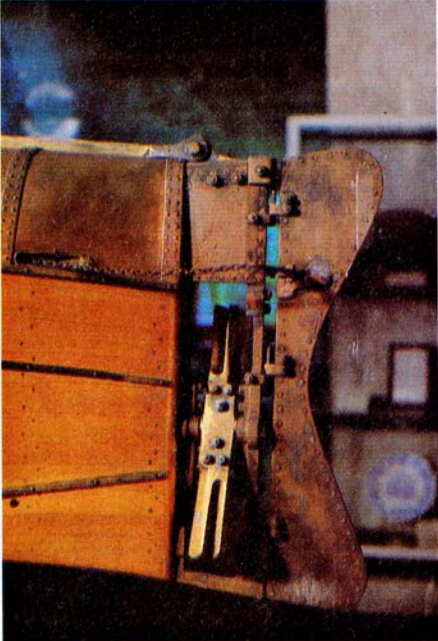
Para la propulsión y viraje se recurría a una hélice, accionada mediante el esfuerzo humano en los primeros ensayos, y al timón. La hélice propulsora, semejante a la del *Ictíneo I*, tenía 1,90 metros de diámetro, un metro de paso y 80 milímetros de entrepalas planas, con lo que se conseguía un resbalamiento del 9 por ciento. Para efectuar los giros necesarios cuando estaba parada, la nave poseía, además, dos hélices laterales colocadas en la parte superior, a popa, inclinadas 45 grados de 0,80 metros de diámetro y 0,20 metros de paso, que giraban a razón de 3 vueltas por segundo, y que podían desconectarse de la transmisión general, alojándose en unos huecos destinados a ello.

Veinte hombres componían la tripulación; de los cuales, 16 estaban inicialmente destinados a servir de motor; la profundidad máxima alcanzada fue de 30 metros.

Uno de los logros más notables del *Ictíneo II* fue la perfección del movimiento vertical de bajada y subida de la nave, es decir, de la inmersión y la emersión. Así, a las vejigas de flotación, examinadas ya al relatar las características del *Ictíneo I,* y que permitían conseguir igualar la densidad de la nave a la del agua, lo cual representaba que con los tanques llenos el submarino permanecía entre dos aguas –con la torreta justo en la superficie–, Monturiol añadió al *Ictíneo II* un sistema que él bautizara con el nombre de vejiga natatoria, encargado de lograr, y dominar, un suave movimiento vertical de la nave. Constaba la "vejiga" de un depósito de agua de lastre (g) comunicado con una bomba de compresión de tres cuerpos (g_2), que a su vez lo estaba con dos depósitos llamados de presión (g_1), sitos en la parte superior de popa detrás de la torreta central, inicialmente llenos de gases y agua, mitad y mitad, a una presión doble de la que correspondía a la profundidad que deseaba alcanzarse [*véase la ilustración de estas dos páginas*]; lo cual representaba un coeficiente o factor de seguridad 2. (Por coeficiente de seguridad se entiende aquí el valor proyectado de una magnitud dividido por valor real máximo alcanzable por de dicha magnitud.)

El funcionamiento del mecanismo era el siguiente: una vez conseguida la posición que hemos llamado de entre dos aguas, es decir, con una densidad





ICTINEO II, visto de frente (*izquierda*) **y en un detalle de su hélice propulsora** (*derecha*). **En su posición frontal, pueden observarse los cinco cristales o ventanas delanteras que permitían la visión y la escotilla de acceso; ésta poseía cuatro cristales más para la observación. Cabe resaltar que la sección del casco exterior perpendicular al eje longitudinal es, prácticamente, una circunferencia. La hélice medía 1,90 metros de diámetro y un metro de paso. Puede apreciarse, a la derecha, el accionamiento del timón mediante cadenas, que durante la navegación permitía girar el navío. (Cortesía del Museo Marítimo de Barcelona.)**

de la nave prácticamente igual o muy poco inferior a la del agua, el timonel accionaba el grifo de entrada de agua al depósito de lastre (g), operación que iniciaba el descenso con una velocidad variable según fuera la cantidad de agua admitida. Cuando se deseaba frenar el movimiento de bajada, anularlo completamente, quedando así en reposo, o bien ascender, bastaba actuar sobre los depósitos g_1 abriendo el grifo que permitía la salida del agua, expulsada por la sobrepresión de los gases, con lo que, al sustituir éstos el volumen que ocupaba aquélla, la densidad media de la nave disminuía. Automáticamente, la bomba g_2 restablecía la situación inicial en los depósitos, preparando una próxima actuación semejante a la descrita, traspasando a los mismos agua del depósito de lastre g; en esta operación la densidad del *Ictíneo* permanecía, pues, constante. De esta manera, un solo hombre dominaba el movimiento vertical del submarino, un movimiento suave y sin sacudidas, por otra parte.

Para llenar los depósitos g_1 –que Monturiol llama vejigas en su Memoria–, no usaba métodos mecánicos, sino que hacía servir los gases productos de la combustión de la hulla por medio de una mezcla de nitrato de sosa y óxido rojo de manganeso (seguramente Mn_3O_4, que se presenta en la naturaleza como un mineral que recibe el nombre de hausmanita); a medida que iba desarrollándose la reacción, el gas obtenido, después de pasar por un depurador que absorbía la sosa cáustica y las posibles pequeñas cantidades de ácido nítrico perjudiciales para los depósitos –por sus efectos corrosivos– llenaba las vejigas hasta alcanzar la presión correspondiente a la profundidad que se quería alcanzar. Así, por ejemplo, si se trataba de preparar un viaje a 50 metros por debajo de la superficie del mar, la presión de los gases tenía que ser de una atmósfera por cada 10 metros, es decir, de cinco atmósferas relativas (o seis absolutas). En tales condiciones, la bomba g_2 introducía agua en ellas hasta doblar la presión inicial: el sistema estaba a punto para efectuar las operaciones antes descritas que originaban el movimiento vertical de la nave.

La seguridad era uno de los factores a tener en cuenta en el *Ictíneo*. Aparte de la sobrepresión del aire encerrado en la vejiga natatoria, el submarino disponía de otros dos sistemas que podían provocar el ascenso en caso de emergencia. El primero de ellos estaba constituido por unas pesas de hierro de forma esférica (K en la sección BB de la ilustración de las páginas 102 y 103); al actuar sobre la manija K', a través de K_1, salían fuera de la nave aligerándola de un peso equivalente al doble del correspondiente al agua encerrada en los depósitos g_1. El segundo de dichos sistemas de seguridad constaba de cuatro cajas de hierro (K en la sección CC) sostenidas por una cadena, de tal manera que, actuando sobre K_1 (en la misma sección CC), las cajas de hierro eran liberadas y se soltaban. En cuanto al equilibrio longitudinal, en el *Ictíneo II* se obtenía fácilmente por medio de un peso desplazable según el eje proa-popa.

La fuerza motriz humana empleada inicialmente en el *Ictíneo II* –en el primer *Ictíneo* no hubo otra–, con todo y corresponder a 16 hombres, sólo fue suficiente para alcanzar una velocidad aproximada de dos nudos (1 metro por segundo) en navegación por la superficie, y de poco más de un nudo (0,5 metros por segundo) en la inmersión, usando una hélice de 1,90 metros de diámetro. Pero la reducida velocidad obtenida, la dificultad e inercia en las maniobras y el incremento del consumo de oxígeno debido al esfuerzo muscular le hicieron pensar en una nueva solución: la aplicación de la máquina de vapor que permitiría doblar –en cuanto a la velocidad superficial se refiere– la alcanzada por fuerza motriz humana.

Ello resultó muy difícil: hubo que construir todos los elementos (caldera, tuberías, condensador, etcétera) por partes para que pudieran pasar a través de la escotilla, cuyo diámetro era sólo de 54 centímetros y ensamblarlas dentro de la nave. Operación, además, cara. Por otro lado, cuando el submarino navegaba por la superficie podía consumir libremente el oxígeno del aire, cosa que no era posible debajo del agua, con lo que se introducía una nueva variable en el ya por sí mismo complicado problema de la navegación submarina, que obligaba a pensar en un generador de oxígeno y en un sistema motriz específico. Sólo se disponía de 29 metros cúbicos para instalar, amén de los aparatos propios, la caldera, dos máquinas de vapor, ventilador para el tipo forzado y purificación de aire, aparato generador de oxígeno, achique, tuberías, pasillos, etcétera.

Para la marcha por la superficie, Monturiol instaló una caldera cilíndrica de 0,8 metros de diámetro y 1,5 metros de longitud (n), en el plano que comunicaba con un depósito de vapor (n_1) de sección elíptica [*véase la ilustración de las páginas 102 y 103*]. Los gases producto de la combustión se evacuaban al exterior a través de una chimenea (n_2), dotada de una válvula de cierre que podía accionarse desde el interior en el momento en que la nave se sumergía. Tanto el horno propiamente dicho como la caja de fuego y los tubos podían desmontarse para su completa limpieza. La caldera era de tiro forzado; el carbón empleado, cok, y la presión, de 4 kilogramos por centímetro

cuadrado. En estas condiciones, la potencia desarrollada era de 6 CV, aunque en las pruebas de determinación de la potencia al freno se había llegado a obtener 8 CV.

La máquina motriz, de la que los planos sólo muestran el cigüeñal (*p*), era bicilíndrica: dos cilindros situados perpendicularmente, con velocidad angular, en régimen normal, de 80 revoluciones por minuto. El condensador (*q*), tubular, estaba situado entre ambos cuerpos de la nave, cubierto normalmente por el agua del mar para facilitar la condensación del vapor saliente de la máquina motriz; una vez condensado, se reciclaba. Se tardaban dos horas, una vez puesta en marcha la caldera, en producir el vapor necesario para accionar el ventilador (*m*). A partir de este momento, la inyección de aire permitía generar la cantidad de vapor necesario, en régimen estacionario, para mover el *Ictíneo* y todos los sistemas adicionales de a bordo.

El problema principal surgía en la inmersión. El cok dejaba de ser el combustible adecuado debido a la producción, en la combustión, de gases nocivos –monóxido y dióxido de carbono (CO y CO_2 respectivamente)–, y a la exigencia de oxígeno que tenía la nave. Consciente de este hecho, Monturiol experimentó numerosas soluciones: las reacciones de nitratos con el hierro y las sulfuraciones de metales. Pero llegó a la conclusión de que la reacción más factible en una cámara herméticamente cerrada, como la del *Ictíneo II*, y que además producía una cantidad de calor apreciable –y suficiente– era la que tenía lugar mezclando clorato de potasa (ClO_3K) con cinc (Zn), en presencia de dióxido de manganeso (MnO_2). Esta reacción tenía, además, una ventaja adicional: en determinada proporción de los reactivos producía oxígeno utilizable para la purificación del aire de la nave.

Para efectuar la mezcla, Monturiol pulverizaba previamente el MnO_2 y a continuación lo calentaba hasta una temperatura comprendida entre los 200 y 300 grados Celsius, con el fin de eliminar el agua que, por su poder higrométrico, llevaba consigo. Luego lo dejaba enfriar lentamente y lo mezclaba con el cinc y el clorato en la siguiente proporción porcentual: cinc, 36 por ciento; dióxido de manganeso, 59 por ciento, y clorato potásico, 5 por ciento. Desde un punto de vista químico la reacción que tiene lugar es bastante compleja, pero podríamos resumirla diciendo que el clorato potásico (ClO_3K) se descompone, al alcanzar una temperatura suficiente, a menudo a través de un estado intermedio, el perclorato potásico (ClO_4K), produciendo finalmente cloruro potásico (ClK) y oxígeno (O_2). En cambio, el dióxido de manganeso (MnO_2) realiza el papel de catalizador, regulando la producción de oxígeno que oxidará al cinc, papel que realiza a través de óxidos superiores, que se descomponen de nuevo en dióxidos de manganeso y oxígeno.

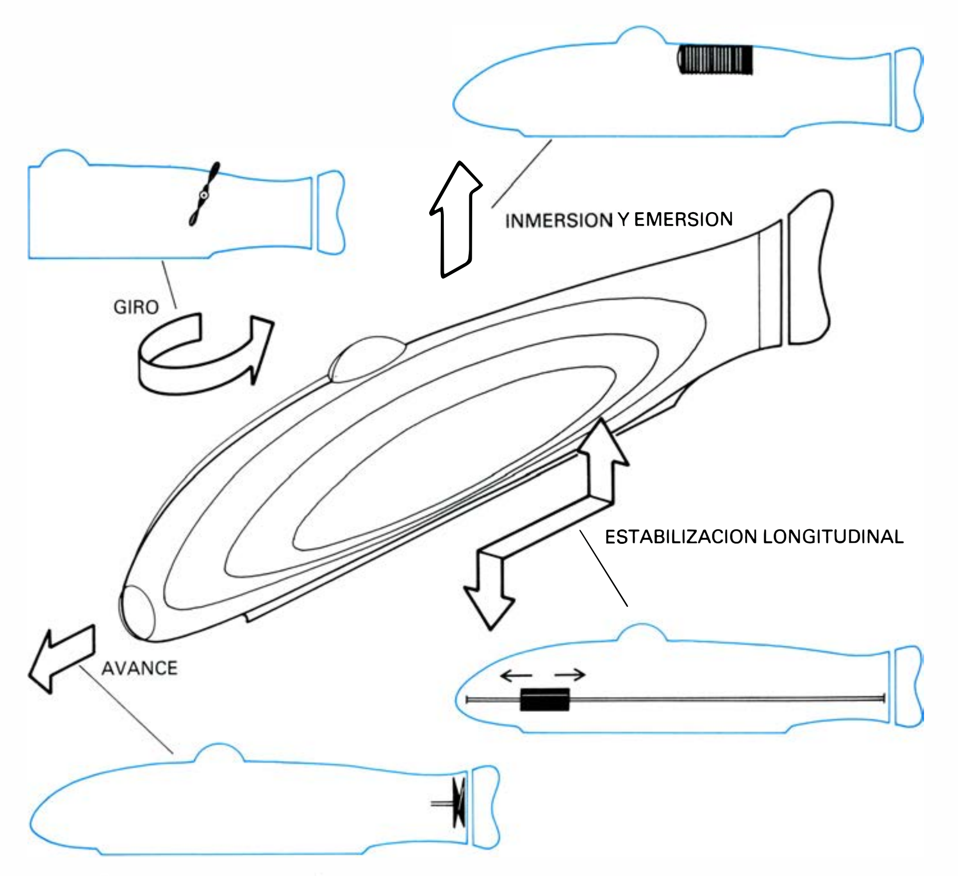
Resumiendo, pues, y prescindiendo de los estados intermedios, digamos que la mezcla de dióxido de manganeso (MnO_2), clorato potásico (ClO_3K) y cinz produce, al reaccionar, cloruro potásico y óxido de cinc (ZnO), reacción exotérmica que libera el suficiente calor para hacerla efectiva. Es evidente, no obstante, que planteaba problemas: no era posible obtener dióxido de manganeso puro (se presentaba mezclado con carbonatos), dando lugar a la formación adicional de anhídrido carbónico (CO_2). Para resolver esta dificultad, tanto el mencionado anhídrido carbónico como el posible oxígeno generado en exceso en la descomposición del clorato potásico se almacenaban en un depósito hasta alcanzar una presión suficiente que les permitía pasar al purificador, donde el segundo era eliminado y el primero, una vez filtrado, podía usarse en la regeneración del ambiente.

La mezcla reactiva se colocaba en unos tubos de 54 milímetros de diámetro exterior, 1 milímetro de espesor y 68 centímetros de longitud, en número de 15, en los que incorporó una "cabeza de ignición" cuya finalidad era, al prenderle fuego, elevar rápidamente la temperatura para que la reacción básica, anteriormente detallada, tuviera lugar con rendimiento apreciable.

Para asegurar la cantidad de oxígeno producida, Monturiol sustituía uno de los tubos o cartuchos, como él les llamaba, por otro donde la mezcla variaba en composición: dióxido de manganeso, 40 por ciento, clorato potásico, 40 por ciento, y limaduras de hierro, 20 por ciento; mezcla que producía, por cada unidad, 300 litros de oxígeno.

El calor liberado por la reacción se empleaba en producir el vapor de agua necesario para accionar una segunda máquina, más pequeña que la que impulsaba la nave por la superficie del mar, monocilíndrica y de eje paralelo a la quilla, que accionaba a su vez un árbol perpendicular al plano de montaje; este movimiento se transmitía al eje longitudinal a través de un par de ruedas helicoidales.

Este procedimiento engendró un



MOVIMIENTOS PRINCIPALES del *Ictíneo II*: el de giro causado por el timón cuando navegaba y las hélices laterales cuando estaba parado; el de avance y retroceso debido a la hélice propulsora; el de cabeceo, que podía corregirse con un peso desplazable a lo largo de un eje longitudinal, y, por último, el de subida y bajada –emersión e inmersión– debido al sistema de vejigas de flotación y natatorias.

nuevo problema: el aumento de temperatura de la cámara habitable. Para solucionarlo, Monturiol razonó que o bien debiera haberse construido la nave con un material buen conductor del calor (la madera es pésimo conductor), cosa que no se hizo por falta de talleres adecuados, o bien refrigerado con agua del mar que circulara a través de tuberías que envolviesen exteriormente el cuerpo interno de la nave. Por qué no lo hizo, nos lo explica él mismo: "Si hubiese construido el segundo Ictíneo con la previsión de un motor fundado en el calórico, la cámara interior no hubiese sido de madera (a pesar de las dificultades que ello hubiera implicado) y hoy nos encontraríamos en las aplicaciones industriales que nos darían los capitales para perfeccionar nuestra obra. Pero los avances en todos los órdenes de ideas son lentos; se presentan sucesivamente, y quien ha de realizarlos depende, generalmente, de personas que no tienen ni pueden tener la perseverancia del inventor, ni su fe en los resultados. Por eso, cuando he tenido necesidad de cubrir la caldera y una parte de las paredes de la cámara con refrigerantes tubulares por donde circulara el agua del mar por diferencia de temperaturas, no he encontrado medios para realizarlo, ni paciencia en los creditores para esperar los resultados de las aplicaciones del Ictíneo".

Tampoco la estancia en el interior del submarino debía ser excesivamente cómoda: una temperatura más alta que la normal en las condiciones de las experiencias, falta de espacio y un ruido infernal (máquinas, ventilador, etcétera) que obligaba a comunicarse a toque de silbato. Pero Monturiol resolvió todos los problemas que se le presentaron con gran ingenio, siendo de justicia citar y constatar la inestimable colaboración de su yerno, José Pascual y Deop, ingeniero industrial y diseñador del complejo motriz.

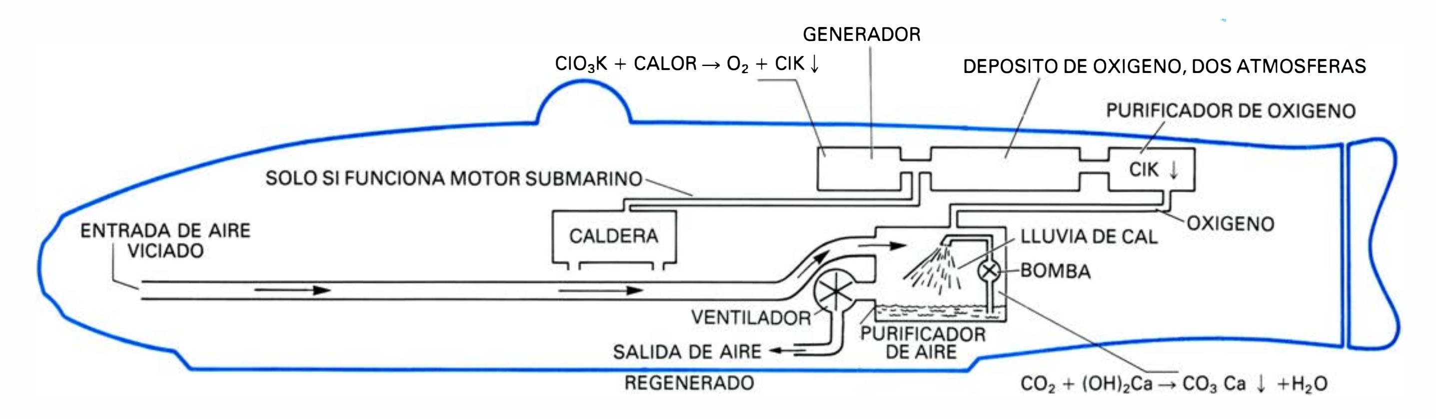
La atmósfera de la cámara habitable le ocupó muchas horas de su tiempo. Y para entender la solución que le diera hay que tener presente el sistema respiratorio del pez, al que él asimilaba su invento. Estableció unos cálculos, siguiendo las teorías de Jean Baptiste A. Dumas, que le llevaron a la conclusión de que se requerían como mínimo 23 litros de oxígeno, por persona y hora, para mantener una respiración correcta en cualquier ambiente, cifra que en las experimentaciones y pruebas realizadas aumentó y concretó en 28 litros. Ahora bien, en realidad la cantidad de aire disponible ha de ser aún mayor; Monturiol, de acuerdo con Jean Claude E. Péclet, la establecía en 6 metros cúbicos por hora y persona: cantidad de aire necesaria para mantener en estado de vapor el agua producida por persona en una hora.

Se realizaron diversas pruebas, en el año 1858, antes de botar el *Ictíneo I*, y posteriormente (1859 y 1860) en la misma nave, aunque todas ellas en condiciones muy diferentes de las establecidas por Péclet, ya que eran de tres a seis los individuos que se encerraban en un recinto de 7 metros cúbicos de capacidad que era, precisamente, la que correspondía a una sola persona según los cálculos citados. Cuando Monturiol incorporó la máquina de vapor al *Ictíneo II*, el problema se agravó: los gases productos de la combustión representaban una nueva variable, no aparecida hasta entonces, que requería y exigía un tratamiento especial. En el primer *Ictíneo*, y debido a la pequeñez de la cámara habitable, se pensó en llevar el oxígeno, embotellado a presión, a bordo; pero esto, en aquella época, era difícil de obtener y, además, caro (según se expuso), por lo que las inmersiones se efectuaron sin ningún tipo de repuesto, es decir, disponiendo sólo del aire inicialmente existente en el submarino y prolongándolas hasta el límite de resistencia de los tripulantes. En cuanto se refiere a la segunda nave, Monturiol abordó el problema frontalmente para llegar a una solución definitiva. Proyectó una instalación de purificación y regeneración de aire muy efectiva: aunque en la navegación por la superficie se podía respirar el aire normal que entraba por la escotilla, la nave disponía de una entrada de aire en la proa (*l*) que podía abrirse y cerrarse a voluntad desde dentro por medio de una válvula plana y de otra análoga en la popa. El ventilador (*m*) era el encargado de mantener una circulación de aire fresco y químicamente puro, tanto si la escotilla estaba abierta como cerrada.

Dichas válvulas se cerraban en el momento de sumergirse. Entraba entonces en acción el sistema de "aire artificial", como le llamaba Monturiol: el ventilador (*m*) aspiraba el aire por un tubo de 20 centímetros de diámetro, en la misma proa de la nave, de manera que los 28 metros cúbicos que correspondían al espacio habitable de la misma eran reciclados, aproximadamente, unas catorce veces por hora, pasando, en cada reciclaje, por un depósito purificador y recibiendo en su trayecto el oxígeno requerido creado en el generador.

En un principio, con fuerza motriz humana, el generador era sólo una cámara donde se producía la descomposición del clorato de potasa en presencia de dióxido de manganeso y limaduras de hierro, de modo parecido al relatado cuando hemos hablado del sistema productor de calor para generar vapor de

PLANO DETALLADO DEL PURIFICADOR DE AIRE (*m_1* en el plano general). *C* es la caja llena de agua alcalina. Una bomba de inyección actúa entre *e* y *f*, impulsando, a través de esta última válvula, la solución alcalina, produciéndose una mezcla íntima con el aire que llega por el tubo concéntrico con la boquilla *f* (una flecha indica el sentido del aire entrante). *V* es el ventilador (*m* en el plano general y en el texto), que aspira el aire por la parte alta de la caja y lo devuelve a la cámara habitada. *G* señala el grifo cuya finalidad es cargar la caja *C*. Por *O* entra el oxígeno puro previamente generado. (Figura 6.ª de la tercera lámina de los planos del *Ictíneo II*, en "Ensayo sobre el arte de navegar por debajo del agua", de Narcís Monturiol. Fondo bibliográfico de la Escuela Técnica de Ingenieros Industriales de Barcelona.)



CUANDO LA NAVE SUMERGIDA se encontraba en reposo, se producía oxígeno en el generador por descomposición del clorato potásico (ClO_3K). El oxígeno (O_2) se almacenaba en un depósito hasta que la presión alcanzaba dos atmósferas, a partir de cuyo momento pasaba a un purificador que retenía el cloruro potásico (ClK). El oxígeno generado y purificado se mezclaba con el aire viciado, en el purificador de aire, donde al mezclarse con agua de cal se producía la eliminación del anhídrido carbónico (CO_2), según la reacción indicada en el esquema. Un ventilador mantenía la circulación de aire reintegrando el aire regenerado otra vez al navío. Cuando funcionaba el motor, en vez de generar O_2 se aprovechaba el sobrante producido en la caldera.

agua en la máquina de vapor introducida por Monturiol en la navegación submarina.

Cuando este procedimiento para obtener la fuerza motriz necesaria se convirtió en realidad, se aprovechó uno de los cilindros de mezcla, como expuse antes, para cargarlo con una proporción de los componentes diferentes de la de los restantes, la cual generaba 100 litros de oxígeno por cada kilogramo de preparación (o 300 litros por cartucho); oxígeno que se sumaba al que podían producir, aunque en menor proporción, los restantes cilindros de carga.

En cualquier caso, el O_2 producido se almacenaba en un depósito hasta alcanzar la presión de dos atmósferas, momento a partir del cual se abría automáticamente una válvula que permitía pasar el gas a una caja con agua, donde se depositaba el cloruro potásico (ClK) presente, en suspensión –producto de la descomposición del clorato potásico (ClO_3K)–. A continuación se introducían el oxígeno generado y el aire reciclado en el purificador; allí la cal apagada –hidróxido cálcico [$Ca(OH)_2$]–, mezclándose íntimamente con ellos, eliminaba el anhídrido carbónico nocivo. (Al ponerse en contacto el hidróxido cálcico y el anhídrido carbónico se produce un precipitado de carbonato cálcico (CO_3Ca) y agua.)

Una vez realizadas estas operaciones de purificación, y con el nuevo oxígeno incorporado, el aire se devolvía a la cámara habitable de la nave. Importa destacar que el generador individual de los tiempos en que eran los hombres quienes movían el *Ictíneo II* se conservó para las operaciones o investigaciones que debía efectuar la nave reposando en el fondo marino.

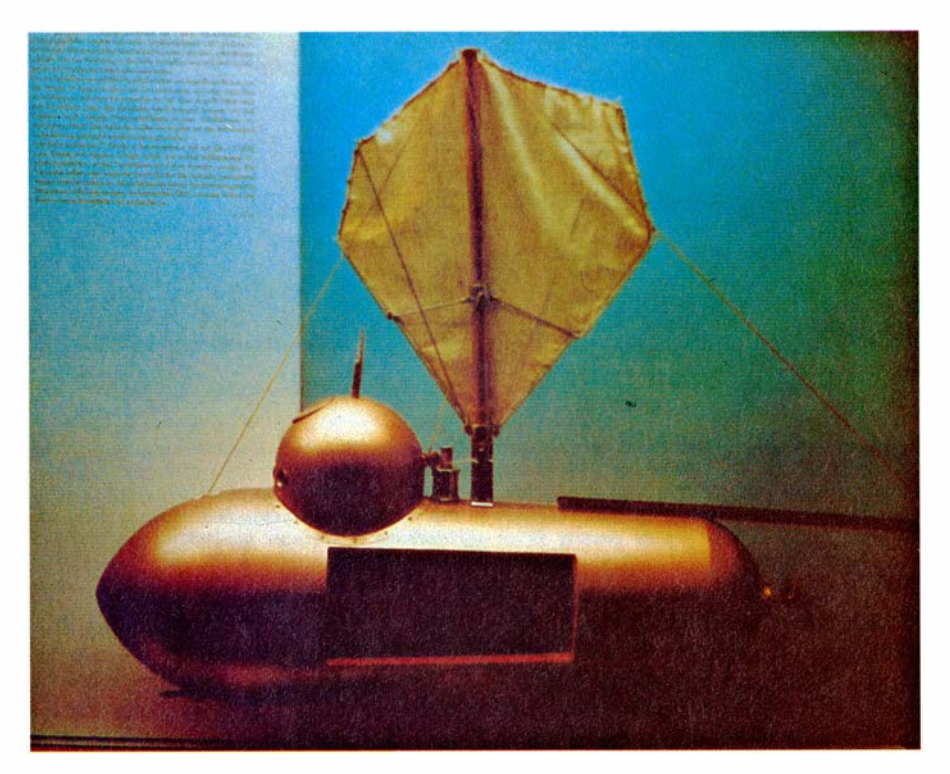
La iluminación interior de las naves se efectuaba con velas. La exterior, que debía ser potente tanto si la nave se empleaba para explorar los fondos marinos como si se pensaba en pesca de coral, se conseguía por medio de la llamada luz de cal, o luz de Drummond, que no es más que la combustión de hidrógeno y oxígeno, en condiciones adecuadas, en presencia de cal viva. El oxígeno se obtenía del que se generaba en la nave; el hidrógeno se llevaba embotellado.

Desde el mes de febrero de 1862 al octubre de 1867 la aventura del *Ictíneo II* no pierde interés. Las dificultades económicas suceden a las técnicas, y si bien éstas son objeto de estudio y resolución por parte de Monturiol, su yerno P. Deop y demás compañeros, aquellas fueron, a la larga, irremontables. "Navegación e industria" amenaza con el embargo de "Navegación submarina" si no procede al pago inmediato de 78.311 reales. Los accionistas, cansados, no responden al último llamamiento y el embargo definitivo se ejecuta el 21 de febrero de 1868, pocos meses antes de la llamada revolución del 68. De nada sirvieron los postreros intentos de D. A. Palés por asumir la deuda con la idea de explotar el invento por su cuenta: la pesca de coral con la nave submarina. El *Ictíneo II* es desguazado y vendido como chatarra.

Los últimos años de la vida del inventor catalán transcurrieron entre el desengaño (por la falta de ayuda que esperó siempre), la desilusión (por no haber podido realizar lo que fue su sueño) y la penuria económica. Había invertido su capital en los *Ictíneos*. No acabaron sus penas: perdió una hija de 17 años y, siete meses más tarde, a un hijo de trece. El año 1873, durante la primera República, fue elegido diputado por Manresa (Barcelona), figurando en el grupo de Castelar. Y prosiguió en su labor creadora: inventó un procedimiento para la fabricación de papel engomado, un original proyecto para llevar las aguas del río Ter a Barcelona, una máquina de hacer cigarrillos, un tranvía funicular, un cañón de montaña de 8 kilos de peso para acabar con los carlistas, un velocípedo, un sistema para conservar la carne... Fracasados todos sus intentos, en posición económica cada vez más deteriorada, murió, en Sant Martí de Provençals, en casa de su hija Anna una mañana de septiembre de 1885. "Muero sin haber podido dar fin a ninguna de mis obras."

Cuando empleamos la palabra invento, normalmente nos referimos a unas aplicaciones prácticas de la ciencia, más o menos complejas, realizadas, la mayoría, dentro de la constante interrelación dialéctica ciencia-técnica. Aunque algunas veces resulta fácil e inmediato unir la realización de una concreta aplicación técnica al nombre de una persona, no siempre sucede así; en muchos casos, esa dialéctica es intensa en su relación y extensa en el tiempo: entonces son varias las personas que contribuyen a su puesta en práctica. En este caso se hace difícil asignar la paternidad de la innovación. Tal ocurre en nuestro caso: el del submarino.

Ahora bien, si esa interrelación ciencia-técnica la representáramos en un diagrama –en el eje horizontal, tiempos; en el vertical, progreso– por una recta inclinada cuya pendiente indicara el grado de avance y la intensidad de perfeccionamiento en la rama cuestionada, aparecerían unos puntos angulosos de discontinuidad de dicha pendiente, que pasaría en ellos de un valor positivo a otro mayor también positivo; esos puntos corresponderían a logros decisivos en el camino emprendido. Si la aportación a la aplicación en cuestión se midiera por el salto de la pendiente, el punto de máxima discontinuidad, en



NAUTILUS, de Robert Fulton. Submarino construido el 1801 gracias al mecenazgo de Napoleón. En la superficie navegaba a vela y sumergido mediante una hélice accionada a mano. (Museo de Munich.)

el caso de la navegación submarina, correspondería a Narcís Monturiol.

En efecto, las naves de Monturiol son de doble casco. Durante mucho tiempo se ha considerado a Labeuf, quien en el año 1899 construyó el "Naval" –un submarino de acero, de 34 metros de longitud y 106 toneladas de desplazamiento–, el inventor del doble casco. Sin menospreciar la labor del ingeniero francés en lo que hace referencia a la navegación submarina, basta contrastar fechas para poder asegurar que Monturiol se había adelantado en este punto a las realizaciones posteriores más destacadas y conocidas. Pero Monturiol no dio ninguna importancia a la construcción con doble casco, pues la encontraba natural: "en lo que respecta al casco del buque, la nave debe cumplir dos condiciones: por un lado, la forma de la construcción debe ser la adecuada a las dos formas de navegación; por otro lado, hay que conseguir que, sumergida, pueda navegar en un medio de presión variable. Para cumplir lo mejor posible estas dos condiciones, he utilizado dos cubiertas, una impermeable y resistente a la presión y otra, exterior, a la que se le ha dado la forma más conveniente para la navegación". Finalmente, con la disposición de doble casco, además de alcanzar una altura metacéntrica adecuada, con una inherente condición de habitabilidad de la nave francamente notable, se disponía de un espacio apropiado para colocar los depósitos de flotación, el condensador y carboneras de cock en el segundo *Ictíneo*.

Otro gran logro del *Ictíneo II* fue el sistema de inmersión, detallado anteriormente y que el mismo inventor le dio el nombre de vejiga natatoria. El sistema empleado de usar gas a presión permitía dominar el ascenso y descenso de la nave mediante la simple manipulación de unas pocas llaves. Este sistema fue adoptado luego por los constructores alemanes.

La solución al problema planteado por el recurso del esfuerzo humano como fuerza motriz (la máquina de vapor) convierte al *Ictíneo II* en un nuevo modelo de nave submarina. A la vez, la resolución de los problemas surgidos en su instalación, así como los que van apareciendo durante la puesta en marcha del sistema, muestran el ingenio de Monturiol, facilitando el camino que habrán de recorrer otros científicos y técnicos que conocerán, así, de antemano, dificultades minuciosamente estudiadas, expuestas y, en su mayor parte, resueltas por él. Concretamente, el sistema que utilizó en la navegación submarina es un hecho demostrativo, no sólo de los conocimientos químicos del inventor catalán, sino también de su creatividad, constituyendo a la vez una de las singularidades que denuncian la creación inventiva.

El sistema de renovación de aire de la cámara habitada es inédito y, además, definitivo. La adición de oxígeno al aire respirado y la eliminación del CO_2 hacen de la nave un lugar que no presenta problemas en este aspecto, por primera vez en la historia. Permiten permanecer bajo el agua durante períodos de tiempo tan largos como se quiera: representa la superación del método del oxígeno comprimido como único medio de subsistir en la cámara cerrada, empleado hasta entonces. Y no se limita a producir oxígeno; también elimina el CO_2. En la conjunción de los dos hechos estriba la pureza de la atmósfera del *Ictíneo*.

Monturiol fue el precursor de las pruebas de seguridad prescritas hoy en todas las naves submarinas. Comprendía perfectamente qué significaba –porque lo había comprobado personalmente– permanecer, vivir con otros compañeros, en un espacio herméticamente cerrado de 29 metros cúbicos; conocía también algunos de los accidentes que habían ocurrido en otros ensayos históricamente anteriores. Concedió una importancia especial al control de todas las variables fisicoquímicas que confluían en las pruebas y a las condiciones de seguridad exigibles. Fue el precursor de la actual "prueba del aire", hoy obligatoria, y que se realiza antes de efectuar cualquier inmersión. Monturiol la ejecutaba sistemáticamente antes de cada prueba, comprobando la estanqueidad de la cámara por medio de las lecturas barométricas leídas y observadas en sendos barómetros colocados en lugares precisos. Haber efectuado más de 50 pruebas en el puerto de Barcelona y cuatro en el de Alicante, sólo con la primera nave, representa superar todos los precedentes y adquirir un certificado de garantía en cuanto se refiere a la seguridad en la navegación submarina: por primera vez en la historia puede hablarse de continuidad.

El espíritu creador e innovador de Narcís Monturiol, así como su formación científica, fruto de un esfuerzo personal elogiable y considerable, le permitieron resolver uno por uno cuantos problemas surgieron en la construcción de las dos naves. Si hubiera vivido –y por tanto hubiera podido desarrollar su labor– unos años más tarde, el consiguiente progreso científico y avance técnico (el uso efectivo de la electricidad) le hubieran permitido clausurar el capítulo de la historia de la técnica dedicado a la navegación submarina.

Juegos matemáticos

Las inspiradas simetrías gráficas de Scott Kim

Martin Gardner

Acaba de salir *Inversions*, de Scott Kim, editado por Byte Publications. *Inversions* es uno de los libros más deliciosos y fascinantes jamás publicados. Con los años, Kim ha perfeccionado hasta un punto increíble su mágica habilidad para tomar una frase no muy larga, o una palabra cualquiera, y rotularla de forma que exhiba alguna asombrosa propiedad de simetría. Fijémonos, por ejemplo, en mi nombre, que vemos más abajo tal como fue rotulado por Kim. Al darle la vuelta, ¡presto! queda exactamente tal cual estaba.

Hace mucho que estudiosos y admiradores de los juegos de palabras descubrieron la posibilidad de construir palabras cortas provistas de diversos tipos de simetrías geométricas. En la Rue Mozart, de París, una tienda de ropa, "New Man", exhibe un gran letrero rotulado "NeW MaN", donde la "e" y la "a" son de formas idénticas, y difieren solamente en su orientación. Todo el rótulo presenta, por consiguiente, simetría central. Los nombres VISTA (revista de la Asociación de las Naciones Unidas), ZOONOOZ (publicación del zoo de San Diego) y NISSIN (un fabricante japonés de equipos de "flash" para fotografía) han sido todos ellos diseñados con gran habilidad, con la intención de que muestren simetría central, es decir, queden invariables al girarlos media vuelta en torno a un punto central.

Hay muchas palabras, como DEDO, DIOXIDO, EXCEDIDO, y algunas frases, como EH, CHICO, HE BEBIDO 33 COCOS, que presentan (casi) simetría respecto de un eje horizontal. Si tras volverlas boca abajo las miramos por reflexión en un espejo situado frente a ellas, quedan invariables. Un día, mi hermana vio intrigada la marca "spep oop" en una caja de galletitas, hasta caer en la cuenta de que se trataba de un envase de "doo dads" que había quedado invertido en el estante. A Wallace Lee, un ilusionista de Carolina del Norte, le encanta ensalzar ante sus amigos las exquisiteces de un plato, "ittaybeds", que él prepara. Para que el nombre no ofrezca duda, lo tiene rotulado en una tarjeta, así:

Ittaybeds

Ante la perplejidad de sus amigos, suele explicar que "saben muchísimo mejor dándole la vuelta al plato".

Muchas palabras cortas, escritas con tipos normales de imprenta, se transforman en otras al invertirlas. Así, NOS se transforma en SON, "up" se convierte en la abreviatura "dn", mientras que "oso", "sosos" y la palabra inglesa SWIMS quedan inalteradas. Otras palabras muestran simetrías respecto de un eje vertical, tales como las inglesas "bid" (y "pig", si se traza la *g* de forma que resulte imagen simétrica de la *p*). He aquí una curiosa forma de escribir "minimum" de modo que no se altere al hacerla girar 180 grados:

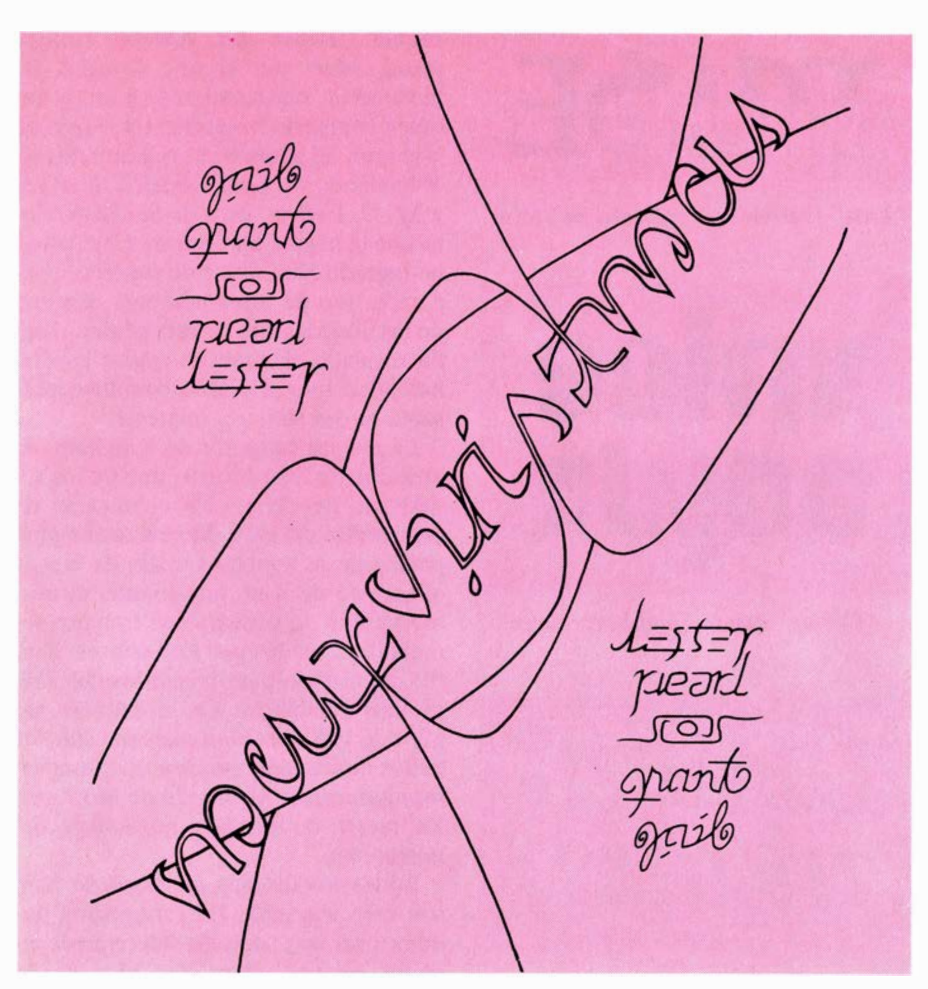
Kim ha logrado elevar este curioso arte de la caligrafía simétrica a cimas que nunca se creyeron alcanzables. Deformando ingeniosamente las letras, aunque nunca tan violentamente que lleguen a resultar irreconocibles, Kim ha conseguido motivos gráficos tan fantásticos como admirables. Su libro es recopilación de estas maravillas, donde engarza llamativas observaciones acerca de la naturaleza de la simetría, sus facetas filosóficas y su integración en las artes plásticas y musicales, lo mismo que en los juegos de palabras.

Kim no es desconocido en esta sección. Hombre joven, nacido en los Estados Unidos, aunque de ascendencia coreana, trabaja actualmente en el departamento de ciencias del cómputo de la Universidad de Stanford. Aún no tenía veinte años cuando empezó a crear problemas sumamente originales de matemáticas recreativas. Entre los publicados en la sección se encuentran los "recorridos de rey perdido" (junio de 1977), el problema de situar caballos de ajedrez en los vértices de un hipercubo (abril de 1978), la solución al "encajado de la caja" (abril de 1979) y su precioso mapa "*m*-perio" (abril de 1980). Además de mostrar sobresaliente capacidad para el pensamiento geométrico (y no sólo en dos o tres dimensiones, sino también en espacios de dimensión cuatro o mayor), Kim es virtuoso pianista, y durante años estuvo dudando si proseguir estudios matemáticos o musicales. En la actualidad, Kim está profundamente interesado en la composición de tipos de imprenta mediante ordenador, campo donde fue pionero su amigo y mentor en Stanford, Donald E. Knuth, distinguida autoridad en ciencias de cómputo.

Durante varios años, el talento de Kim para rotular palabras y dotarlas de las más inesperadas formas de simetría permaneció confinado en el diseño de tarjetas de felicitación navideña para su familia, y como diversión para entretener a su amigos. No era raro que al conocer a alguien en una reunión o una fiesta, Kim se esfumara un ratito, para reaparecer con su nombre pulcramente dibujado en una tarjeta, con perfecta simetría central. En una de las ilustraciones de la página adyacente podemos ver su tarjeta de felicitación de las Navidades de 1977, que presenta simetría central. (Lester y Pearl son sus padres; Grant y Gail; su hermano y hermana, respectivamente.) Al año siguiente, Kim descubrió la forma de que "Merry Christmas, 1978" quedase simétricamente centrado respecto de un eje horizontal y, en 1979, que el eje de simetría fuese vertical.

Para un aniversario de boda de sus padres, Kim ideó un pastel cubierto de vainilla y chocolate, cuyo diseño vemos en la página contigua. ("Lester" se ha trazado en negro, "Pearl", en blanco, está invertido.) Tal es la técnica de "fondo y figura" de Kim. Puede verse otro ejemplo de esta técnica en *Gödel,*

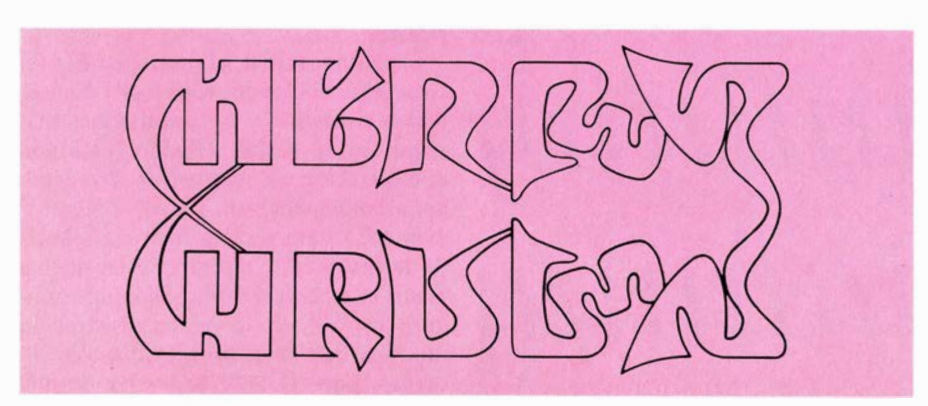
El nombre del autor queda invariable al ponerlo cabeza abajo. ¡Compruébelo!



La tarjeta navideña de Kim se lee también cabeza abajo



"Merry Christmas" tiene simetría respecto de un eje vertical



"Merry Christmas" es ahora simétrico respecto de un eje horizontal



"Lester" y "Pearl" en fondo y figura



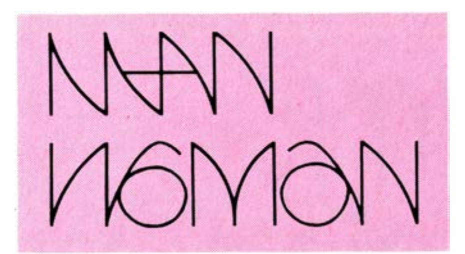
Gödel, Escher y Bach dotados de simetría axial

"True" (verdadero) inmerso en "false"

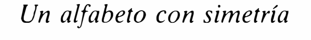
"Upside" queda igual boca abajo



Una "communication" entre dos puede invertirse



"Man" y "woman" pueden invertir sus posiciones



Un alfabeto con simetría

"Infinity" prosigue hasta el infinito en ambos sentidos

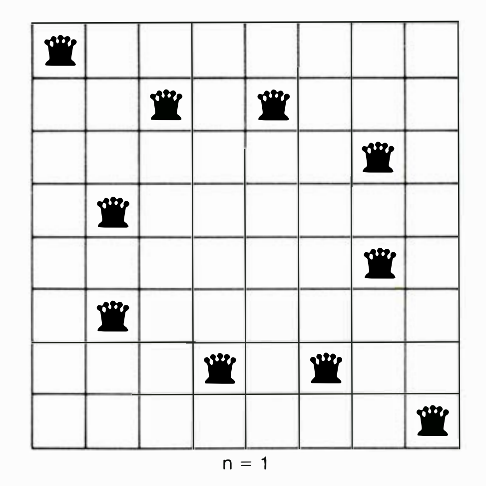
Escher, Bach: An Eternal Golden Braid, libro con el que Douglas R. Hofstadter, buen amigo de Kim, y con quien comparto la sección a lo largo de este año, ha ganado el premio Pulitzer. Y hablando de Kurt Gödel, J. S. Bach y M. C. Escher, otra de las ilustraciones de la página anterior muestra cómo ha logrado Kim dotar de simetría axial a cada uno de estos nombres. En una de las ilustraciones de esta página, Kim ha rotulado el alfabeto entero con tal habilidad que el diseño completo presenta airosa simetría bilateral.

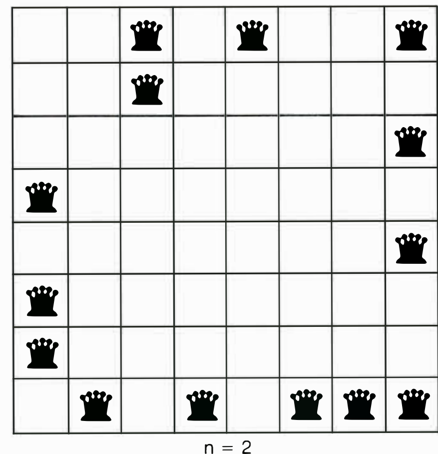
La mágica caligrafía de Kim llamó la atención de Scot Morris, uno de los redactores de *Omni*. En el número de septiembre de 1979, Morris dedicó una página de su popular sección de juegos al trabajo de Kim, anunciando al mismo tiempo un concurso de trabajos semejantes creados por los lectores. Kim fue contratado para juzgar los miles de diseños recibidos. En el número de abril de 1980 de *Omni* pueden verse los bellos motivos premiados; otros finalistas aparecen en la sección de Morris en los meses de mayo y noviembre del mismo año.

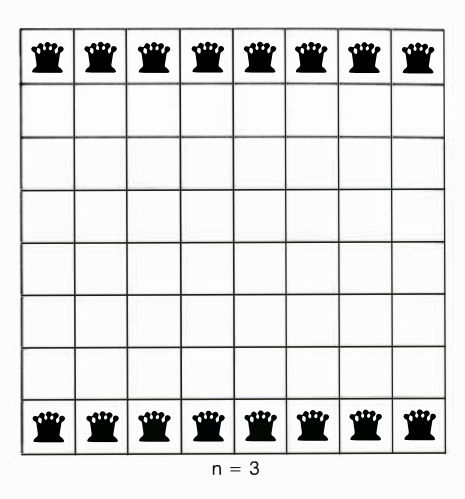
Todos los diseños del libro de Kim son creación suya. En esta página podemos ver una pequeña selección de algunos nuevos, para dar idea de la asombrosa variedad de recursos visuales que Kim aún guarda en la manga.

Dedicaré ahora una mirada a dos insólitos problemas matemáticos, sólo parcialmente resueltos, originales de Kim. En 1975, estando Kim todavía en secundaria, se le ocurrió una notable generalización del clásico problema de situar ocho reinas sobre un tablero de ajedrez sin que ninguna amenace a ninguna otra. Preguntémonos, dijo Kim, cuál es el número máximo de reinas que podemos colocar en el tablero, con la condición de que cada una amenace a exactamente otras n, siendo n un número dado de antemano. Las reglas ordinarias de ajedrez siguen en vigor, así que ninguna dama podrá amenazar a otra situada tras una tercera interpuesta.

Cuando n vale 0, el problema es precisamente el clásico. Kim logró demostrar que cuando $n = 1$, el número máximo era 10 reinas. (Puede verse una demostración en *Journal of Recreational Mathematics*, vol. 13, n.º 1, pág. 61; 1980-81.) Vemos en la figura izquierda de la ilustración superior de la página contigua una muy bella solución. La figura central de la misma ilustración muestra una solución máxima, con 14 damas, para el caso de $n = 2$; en una carta, donde comentaba tal disposición de las piezas, Kim decía que "es tan







Problemas de las reinas en el tablero, de Kim

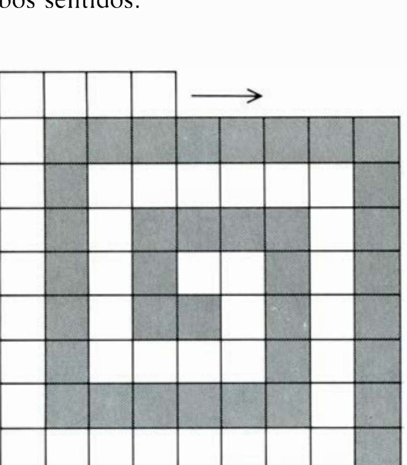
horriblemente asimétrica que no tiene derecho a existir". Para los casos $n = 3$ y $n = 4$, sólo hay conjeturas en cuanto al número de reinas del problema. El mejor de los resultados conseguidos por Kim para el caso $n = 3$ consta de 16 damas, alineadas de forma ridículamente simple. Cuando n es 4, lo mejor que ha conseguido Kim es situar 20 reinas. ¿Conseguirá el lector hacer otro tanto, esto es, situar 20 damas sobre un tablero de ajedrez, de manera que cada una amenace a otras cuatro?

Evidentemente, el problema puede generalizarse a tableros de dimensión arbitraria; pero Kim ha demostrado, valiéndose de teoría de grafos, que en ningún tablero, por grande que sea, podrá existir solución cuando se exija que n sea mayor que 4. Para el caso $n = 1$, Kim ha logrado una demostración sencilla de que el número de reinas no puede superar al mayor de los enteros menores o iguales que $4k/3$, siendo k el número de cuadros del lado del tablero. Para $n = 2$, Kim dispone también de una demostración, aunque no tan sencilla, que establece que el máximo número de reinas no puede superar $2k - 2$ e, igualmente, que dicho límite es alcanzable en todos los tableros de orden par.

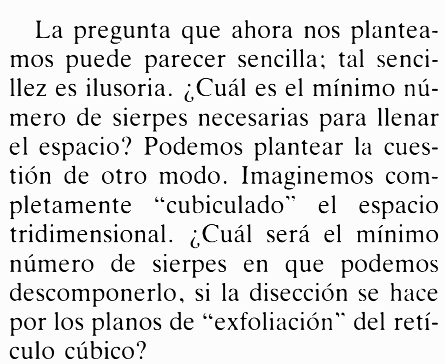
El problema de Kim sobre sierpes policúbicas no se ha publicado con anterioridad, y tanto él como yo recibiremos con agrado e interés la luz que sobre el mismo puedan arrojar los lectores. Definamos ante todo qué son las sierpes. Las sierpes son cadenas conexas de cubos unitarios idénticos, que se tocan por sus caras, de manera tal que cada cubo (excepto los situados en cabeza y cola, si los hay) está en contacto, cara a cara, con otros dos (y sólo dos) cubos de la cadena. La sierpe puede torcerse y retorcerse en todas las direcciones posibles, con tal de que se respete la regla de contacto explicada. No hay objeción, empero, a que al retorcerse la sierpe queden en contacto a lo largo de sus aristas o a través de sus vértices un número cualquiera de cubos. Hay varios tipos de sierpes policúbicas. Pueden ser abiertas, de longitud finita, con dos cubos extremos –a modo de cabeza y cola– que sólo están en contacto con un cubo, y no con dos. Pueden ser cerradas, de longitud finita, sin extremos, a modo de collar. Las infinitas pueden tener origen o no; en este último caso serán ilimitadas en ambos sentidos.

La pregunta que ahora nos planteamos puede parecer sencilla; tal sencillez es ilusoria. ¿Cuál es el mínimo número de sierpes necesarias para llenar el espacio? Podemos plantear la cuestión de otro modo. Imaginemos completamente "cubiculado" el espacio tridimensional. ¿Cuál será el mínimo número de sierpes en que podemos descomponerlo, si la disección se hace por los planos de "exfoliación" del retículo cúbico?

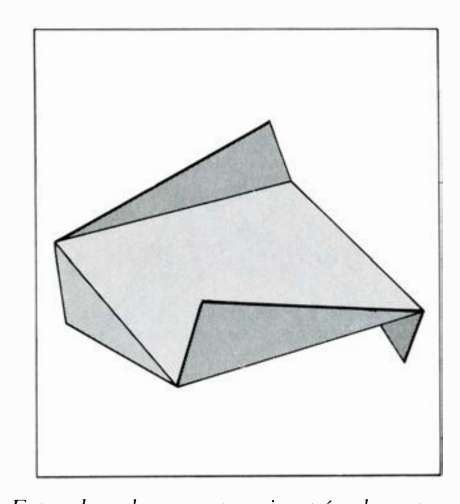
Al estudiar el homólogo bidimensional de este problema (ahora las sierpes

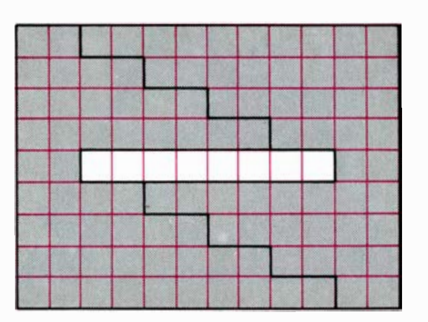


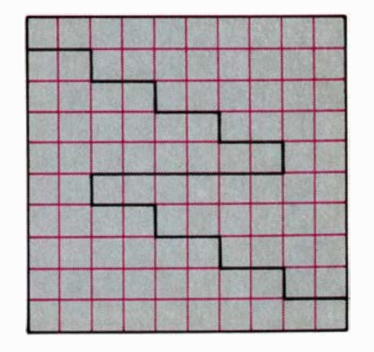
Dos sierpes teselan el plano



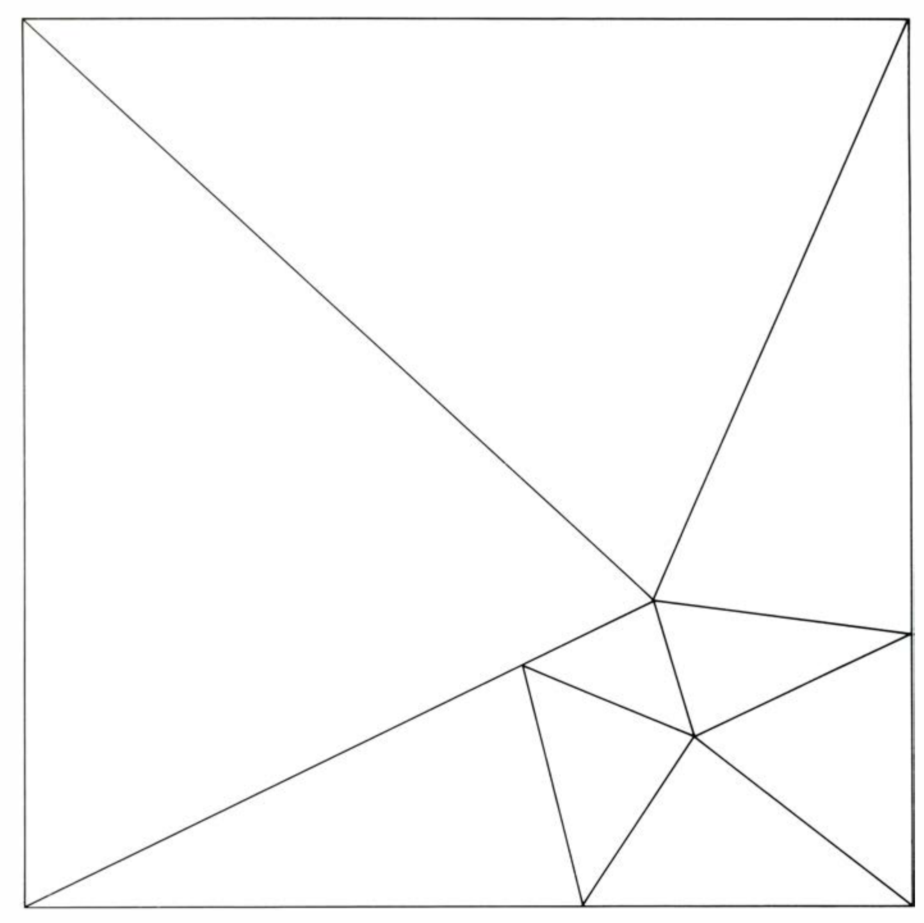
Este plegado muestra simetría de rotación especular



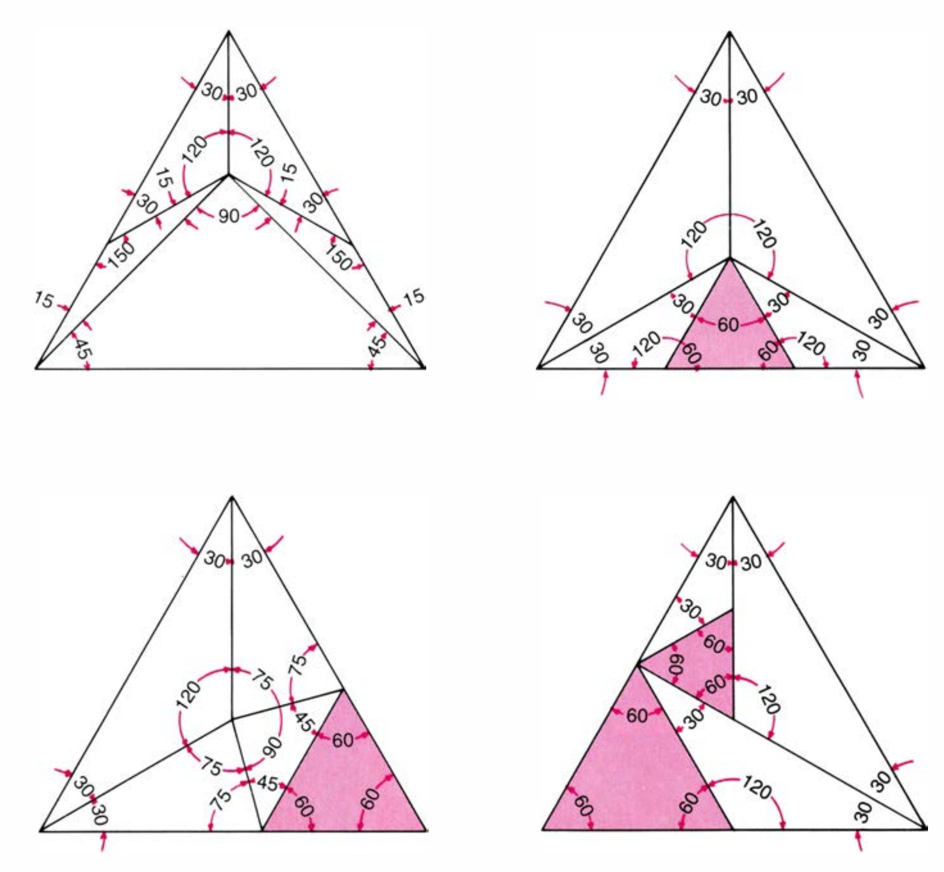




Una colcha de retazos vuelta a coser para lograr otra cuadrada



Descomposición de un cuadrado en nueve triángulos acutángulos



Cuatro disecciones de un triángulo equilátero en cinco triángulos isósceles

estarán formadas por cuadrados unitarios) es sencillo ver que la solución es dos. Basta para ello entrelazar dos sierpes espirales infinitas, que parten de orígenes contiguos, una blanca, y otra gris, como vemos en la figura superior izquierda de la ilustración inferior de la página anterior.

Cómo llenar el espacio con sierpes policúbicas ya es harina de otro costal. Kim ha descubierto un método para formar con cuatro sierpes que parten de otros tantos cubos un "ovillo" –conviene imaginar que cada sierpe es de un color– formado por sierpes helicoidales entretejidas, que llena todo el espacio. Su método es demasiado complejo para tener cabida en este reducido espacio; los lectores tendrán que confiar en mi palabra: puede lograrse.

¿Podrá lograrse con sólo tres sierpes? No sólo permanece abierta esta cuestión, ¡Kim ha sido incapaz de demostrar que es imposible hacerlo con dos! En una carta, Kim decía: "Una solución con sólo dos sierpes crearía una especie de símbolo yin-yang tridimensional e infinito, donde cada sierpe sería el espacio que deja libre la otra. Tal sería la belleza del entrelazado, y tan atrayente la idea de construir un modelo lo bastante grande como para reptar por su interior, que persisto en buscar una solución".

Evidentemente, el problema puede generalizarse a sierpes formadas por cubos unitarios de cualquier número de dimensiones. Kim ha conjeturado que, en un espacio de n dimensiones, el mínimo número de sierpes que lo llenarán completamente será $2(n - 1)$, pero tal conjetura no tiene todavía bases muy firmes.

Hace algunos años tuve el placer de explicarle el problema de las sierpes policúbicas a John Horton Conway, famoso matemático de Cambridge. Al terminar le dije que Kim no había demostrado todavía la insuficiencia de dos sierpes para rellenar el espacio tridimensional. Conway replicó instantáneamente, "Pero es evidente que...". Se contuvo a mitad de la frase, quedó mirando al espacio uno o dos minutos, y después exclamó "¡*No* es evidente!"

No tengo la menor idea de los pensamientos que cruzaron la mente de Conway. Sólo puedo decir que si ni Kim ni Conway consideraron evidente la imposibilidad de llenar el espacio con dos sierpes, probablemente tampoco le resulte evidente al resto de los mortales.

He aquí las soluciones a los problemas propuestos en junio:

1. Fijémonos en los dos puntos extremos, A y K. Todos los puntos L si-

tuados sobre el segmento que une A con K (ambos inclusive) tendrán la misma suma de distancias a A y a K. Fijémonos ahora en B y J, que es el siguiente par de puntos al movernos hacia el interior. Como antes, para hacer mínima la suma de distancias de L a B y a J, L deberá encontrarse entre B y J. Como L se encuentra también entre A y K, tal situación hará mínima la suma de las distancias desde L hasta A, K, B y J.

Prosigamos de igual forma, tomando los puntos por pares conforme vamos introduciéndonos más y más en el encaje de intervalos. El último par de puntos es E y G. Entre ellos queda el punto F. Cualquier lugar elegido entre E y G hará mínima la suma de distancias a todos los puntos, salvo a F. Evidentemente, para hacer también mínima la suma de distancias a F, el lugar elegido debe estar precisamente en F. En pocas palabras, Lavinia tendría que irse a vivir al mismo edificio donde se aloja su amigo Frank.

Generalicemos. Tomando sobre una recta un número impar de puntos, todo emplazamiento elegido entre los dos puntos más interiores será solución del problema. Cuando el número de puntos sea impar, la solución (ahora única) corresponde al punto "central". El problema se publicó en "No Calculus, Please", de J. H. Butchart y Leo Moser (*Scripta Mathematica*, vol. 18, n.^{os} 3-4 págs. 221-236; 1952).

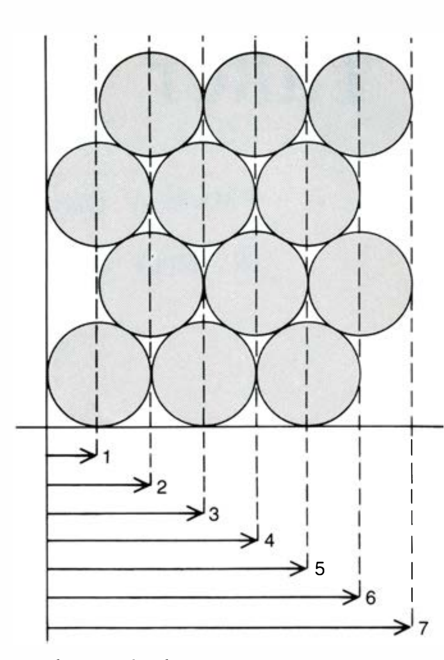
2. La figura superior derecha de la ilustración inferior de la página 113 muestra cómo plegar un cuadrado de papel y formar una figura carente de planos de simetría, que puede en cambio superponerse a su imagen reflejada en un espejo. Esta figura posee la llamada "simetría de rotación especular", tipo de simetría de gran importancia en cristalografía. He tomado el ejemplo de la página 42 de *Simmetry in Science and Art*, por A. V. Shubnikov y V. A. Koptsik (Plenum Press, 1974).

3. Las figuras inferiores de aquella ilustración muestran cómo reparar la colcha de retazos deteriorada. La línea quebrada en negro indica por dónde deberemos descoserla para que las dos piezas resultantes, ensambladas, den un cobertor cuadrado, sin agujeros. El problema figura con el n.º 215 de los *Puzzles and Curious Problems*, de Henry Ernest Dudeney (Thomas Nelson and Sons, Ltd., 1931).

4. En la figura superior de la página precedente podemos ver una descomposición de un cuadrado en nueve triángulos acutángulos. Probablemente la solución sea única. Tomando la triangulación en sentido topológico, será inadmisible que haya vértices situados en lados, y se deduce de aquí la imposibilidad de descomposiciones con nueve triángulos, si bien las hay con ocho, 10 y todos los valores posteriores. Este curioso resultado ha sido establecido por Charles Cassidy y Graham Lord, de la Universidad de Laval.

La ilustración inferior muestra cuatro disecciones de un triángulo equilátero en cinco isósceles. En la primera, ninguno de estos cinco triángulos es equilátero. La segunda y tercera contienen cada una un equilátero y, la cuarta, dos. Las cuatro descomposiciones, ideadas por Robert S. Johnson, aparecen en *Crux Mathematicorum* (vol. 4, n.º 2, pág. 53; febrero, 1978). En el mismo volumen de esa revista podemos ver una demostración de Harry L. Nelson sobre la imposibilidad de que la disección contenga más de dos triángulos equiláteros (n.º 4, págs. 102-104; abril, 1978).

5. La ilustración superior de esta página muestra cómo podemos medir distancias de número entero de centímetros valiéndonos de monedas de 1 yen, que tienen exactamente un centímetro de radio.



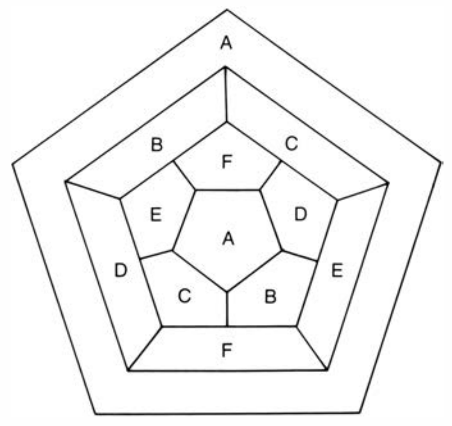
Medición de distancias enteras con monedas de un yen

6. El mapa de la ilustración inferior, descubierto por Lloyd Shapley, matemático de la Rand Corporation, demuestra que cuando se usan cinco colores en una partida de coloreado de mapas (inventada por Steven J. Brams) existe un mapa con el cual el maximizador logra siempre vencer.

El mapa es proyección del esqueleto de un dodecaedro, con una región exterior (A) que representa la cara "trasera" de este poliedro. La estrategia del maximizador consiste en colorear la cara del dodecaedro diametralmente opuesta a la última de las iluminadas por su adversario, usando precisamente el mismo color. (En la ilustración, las regiones que representan pares de caras opuestas están señaladas con la misma letra.) Como fácilmente podemos comprender, tal estrategia va eliminando sucesivamente los colores utilizados, impidiendo su empleo posterior. Ello lleva el juego a una situación donde el minimizador no puede actuar sin servirse de un sexto color.

¿Existirá un mapa donde el maximizador pueda obligar a su contrario a emplear siete colores, aunque éste realice el mejor de sus juegos posibles? Tal problema permanece aún sin resolver.

7. La estrategia de John Horton Conway para su juego "whim" es como sigue. La "jugada de capricho" recibe el mismo tratamiento que una hipotética pila, formada por una ficha en el caso de que haya algún montón compuesto por cuatro o más cuentas, y formada por dos cuentas si ninguna pila contiene al menos cuatro fichas. Mientras ningún jugador haga uso de la jugada de capricho, la hipotética pila permanece en su (invisible) sitio. Ejercer el derecho de capricho equivale a retirar la pila fantasma. En una partida de whim, cada posición del juego dará la victoria al jugador que llevaría las de ganar en el *nim* ordinario si la pila imaginaria estuviera materializada sobre el tablero. Por consiguiente, la estrategia vencedora consiste nada más en imaginar que la pila invisible se encuentra verdaderamente en la mesa, hasta que uno de los jugadores la retire, y jugar siempre la estrategia vencedora en la partida de *nim* ordinario. Cuando una de las jugadas cambie la posición del juego, pasando de una donde al menos una pila tenía al menos cuatro piezas a otra dónde las pilas tienen menos de cuatro, la pila "de capricho" recibe su segunda cuenta tras efectuarse este movimiento, no antes.



Una demostración de cómo vencer en el coloreado de mapas

Taller y laboratorio

Física y química subyacentes a la tarta de merengue al limón

Jearl Walker

Los libros de cocina no suelen detenerse en los aspectos científicos del arte del cocinar; no consideran las transformaciones que sufren los alimentos cuando se guisan. Una escaramuza en ese campo me llevó hasta los misterios de la salsa bearnesa, que ya les describí en febrero de 1980. Hoy le toca al merengue. ¿En qué procesos se ven involucradas las claras de huevo cuando las batimos y calentamos? Quizá mi menguado éxito espolee a otros a experimentar en este terreno.

Un plato de merengue plantea algunas cuestiones sobre la química de la proteína de los huevos y la física de la transmisión de calor. Por ejemplo, ¿por qué se espesan los huevos cuando se baten? ¿Por qué un batido excesivo separa agua de las claras y estropea un postre hecho con ellas? ¿Por qué la adición de un ácido débil, como vinagre, da consistencia a las claras? ¿Por qué se prohíbe la adición de sal antes de batir las claras?

Sea nuestro objetivo una tarta de merengue al limón. El merengue es, en ella, un copete esponjado formado por claras de huevo batidas y cocidas al horno a las que se ha añadido una pizca de azúcar. ¿Cómo hacerlo? Se empieza por separar las claras de las yemas. Se cascan los huevos y se vierte su contenido en un separador de huevos, en el que la depresión central retiene la yema y la clara se desliza por el reborde del separador para caer en un cuenco. Bajo ningún concepto deben entrar las yemas en el merengue; la mejor manera de asegurarse de que así sea consiste en separar, en un pequeño cuenco, la clara de cada huevo. Convencidos de que no pasó yema al cuenco, podrá añadirse el contenido de éste a otro cuenco mayor, donde están las claras separadas hasta entonces.

Se baten luego las claras con un batidor de alambre, una paleta de cocina o una batidora eléctrica. Esa operación no tiene más finalidad que espesarlas e introducir en ellas burbujas de aire poco a poco. Se seguirá batiendo hasta que la sustancia se torne opaca, reluciente y consistente. En este punto puede de introducirse el azúcar, batiéndolo paulatinamente; la mezcla puede depositarse entonces, con cuidado, sobre la base de limón. Ya puede entrar la tarta en el horno.

El calor de la cocción dilata y endurece el copete de claras. La dilatación resulta del aumento de tamaño que sufren las burbujas de aire atrapado cuando se calientan; el endurecimiento, de la cocción de los huevos. El merengue es liviano y mullido porque el aire que se le introduce por batido rebaja su densidad, que decrecerá aún más cuando el aire se dilate en el horno.

¿Por qué han de estar las claras limpias de yema? Por la sencilla razón de que, aun cuando la cantidad fuera muy pequeña, la operación de batido no añadiría aire suficiente, y resultaría un merengue denso y desinflado. El dulce puede frustrarse también por exceso de batido. Así, en mi afán por espesar las claras, les doy muchas veces más vueltas de la cuenta, provocando que el agua se separe; este error se paga con un merengue desinflado y duro.

Hace poco, y a raíz de otro fracaso con ese pastel, me planteé con resolución qué es lo que en realidad les ocurría a las claras de huevo durante el batido. ¿Cuál es la causa de que se transformen, desde un estado de fluidez transparente, en una mezcla opaca de huevo y burbujas de aire? ¿Y qué le pasaba a la mezcla cuando la batía más de lo debido?

Estas preguntas me llevaron a cavilar acerca de otros procedimientos. Ciertos cocineros añaden una pequeña cantidad de ácido, ya sea vinagre (ácido acético) o crémor (ácido tartárico), para favorecer el espesamiento de las claras; la sal, por el contrario, se desecha ya que las endurece, según se dice. No debe añadirse azúcar hasta que las claras se hayan batido por completo y entonces se verterá poco a poco, batiendo ligeramente más, aunque lo menos posible. Por último, hay cocineros que afirman que el mejor merengue se consigue batiendo las claras en un cuenco de cobre.

Para explicarles qué les ocurre a las claras siguiendo estos procedimientos voy a describirles primero la naturaleza proteínica de la clara. Los buenos cocineros preparan esas proteínas para el horno; con mi escasa pericia, lo que hago yo es estropearlas. El éxito de un merengue depende de cómo se comporten las proteínas cuando se baten y se combinan con ácidos débiles.

Las claras de huevos de gallina constan de varios tipos de proteínas, principalmente ovalbúmina, globulina y ovomucina. Las proteínas son moléculas gigantes formadas por una cadena (o varias) de aminoácidos unidos por enlaces peptídicos. Estos son enlaces covalentes fuertes, tendidos entre un átomo de carbono de un aminoácido y un átomo de hidrógeno del aminoácido siguiente de la cadena. En la ilustración superior de dos páginas más adelante se representa una cadena proteínica genérica, en la cual el símbolo R representa la cadena lateral que caracteriza a cada aminoácido. La ovalbúmina se compone de 386 aminoácidos.

La secuencia de aminoácidos de una cadena proteínica constituye la estructura primaria de la proteína. Dicha cadena no es recta, sino que posee una estructura secundaria tridimensional. En algunos tipos de proteínas, la cadena forma una hélice; en otras, se asemeja a una hoja de papel plegada. La ovalbúmina y la globulina pueden presentar una mezcla de ambas junto con regiones en donde las cadenas se arrollan más al azar. Estas estructuras secundarias pueden replegarse para formar una estructura terciaria. Así, la ovalbúmina y la globulina se repliegan de modo tan complicado que parecen esferas y no cadenas; las proteínas que presentan esta configuración se denominan globulares.

Las estructuras secundaria y terciaria se mantienen merced a la acción de fuerzas de distinta naturaleza. Las uniones bisulfurosas son enlaces covalentes entre átomos de azufre situados en puntos distintos de una misma cadena. Otros puntos de esta pueden unirse a través de enlaces iónicos, que se originan por atracción electrostática entre dos lugares con cargas de signo contrario.

Los lugares neutros (o sea, sin carga) pueden atraerse entre sí a través de una fuerza que en biología se llama enlace hidrófobo y, en física, fuerza de van der Waals. En este caso acontece lo siguiente: desde uno de los lugares se fuerza una pequeña separación entre los centros de las cargas negativas y positivas del otro, y esta separación hace que las cargas positivas que hay en un lugar se vean atraídas por las cargas ne-

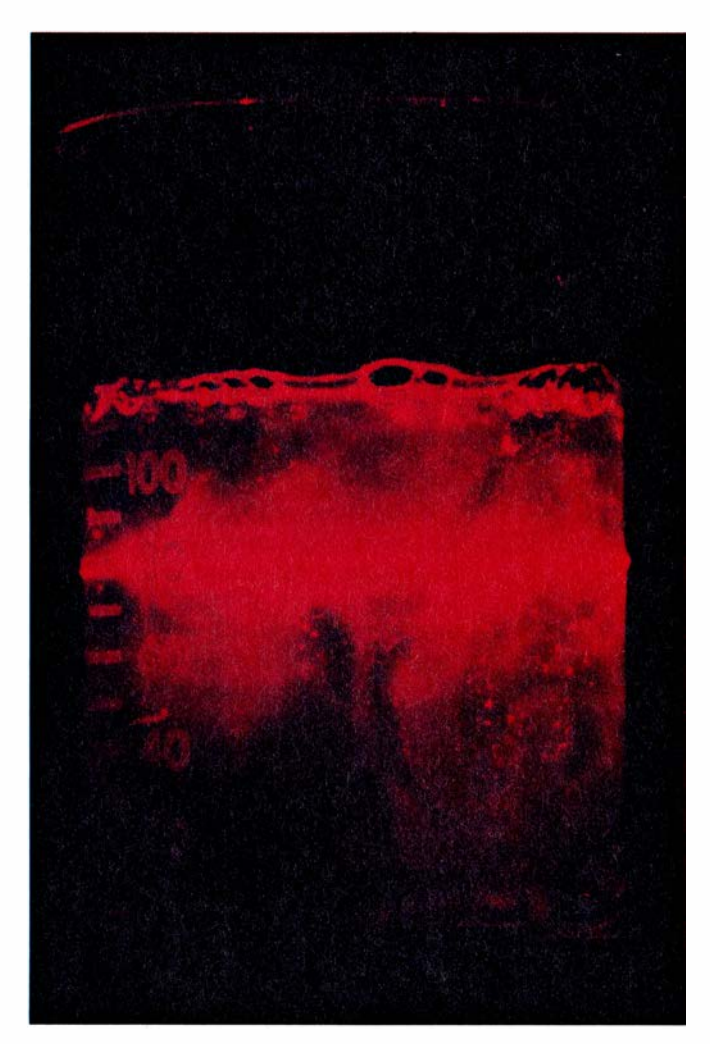
gativas del otro. En los grupos polares sucede algo similar, aunque en ellos es permanente la separación entre los centros de las cargas positivas y negativas. El extremo positivo de un grupo polar es atraído entonces por el extremo negativo del otro grupo. Corrientemente, el extremo positivo es un átomo de hidrógeno, en cuyo caso la atracción se llama enlace hidrogénico.

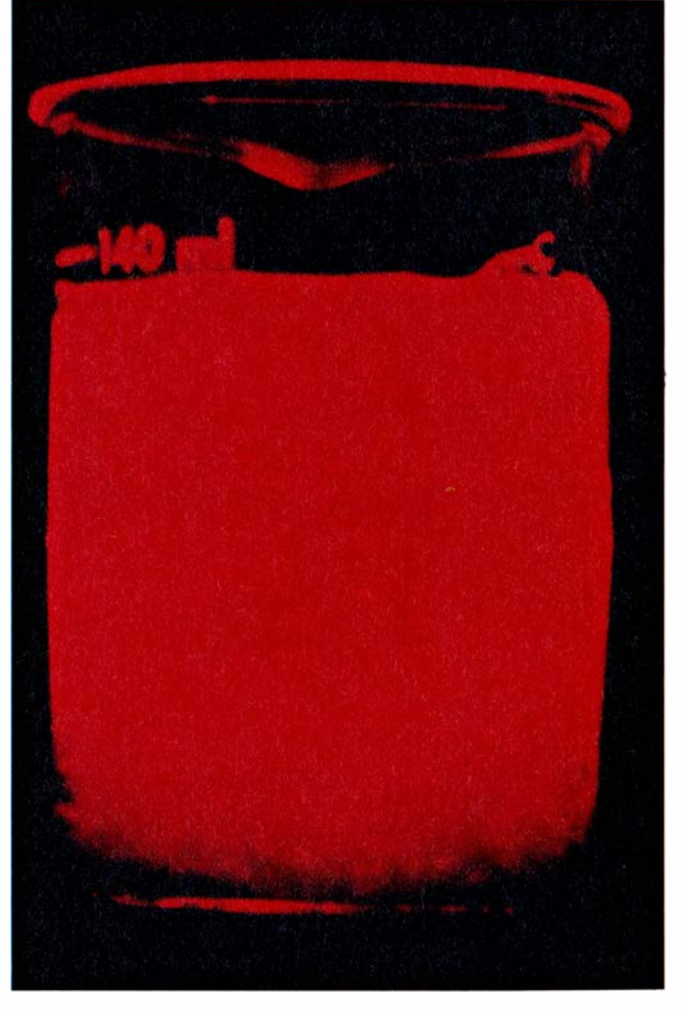
De todas estas interacciones, las más intensas son las peptídica y la bisulfurosa, ambas covalentes; siguen, a gran distancia, los enlaces iónicos. Los enlaces hidrófobos y los hidrogénicos son mucho más débiles todavía. Cuando un cocinero introduce un batidor de alambre dentro de las claras, hendiendo el fluido, parte de los enlaces más débiles se rompen y se destruyen algunas porciones de la estructura terciaria. Pero las proteínas no se despliegan por completo, ya que las fuerzas que mantienen las estructuras primaria y secundaria son relativamente intensas. Como consecuencia del batido, las proteínas se desenredan poco a poco, a partir de un estado inicial de spaghetti apelotonado. Toda alteración de una estructura proteínica realizada por este procedimiento se llama desnaturalización. Así pues, uno de los objetivos que se persiguen al batir huevos es desnaturalizar las proteínas.

Desenmarañadas en parte, las moléculas proteínicas comienzan a unirse entre sí para formar una malla tridimensional, o gel. Esta unión entre las proteínas resulta improbable antes de la desnaturalización, puesto que las proteínas son globulares y exponen relativamente pocos puntos propicios para establecer enlaces. Además, los lugares más adecuados para enlazarse hidrofóbicamente están escondidos, en un comienzo, en el interior de las proteínas globulares. Ahora bien, cuando se desenredan en parte aparecen más lugares por donde unirse a otras. A medida que avanza el batido se forman burbujas de aire (aire rodeado de una película de agua delgada), que quedan retenidas en la malla de gel. La película de agua mantiene su cohesión gracias a la tensión superficial y se ve reforzada por la malla de proteínas que contiene. Además de unirse entre sí, estas proteínas se enlazan (hidrogénicamente) a las moléculas de agua de la película.

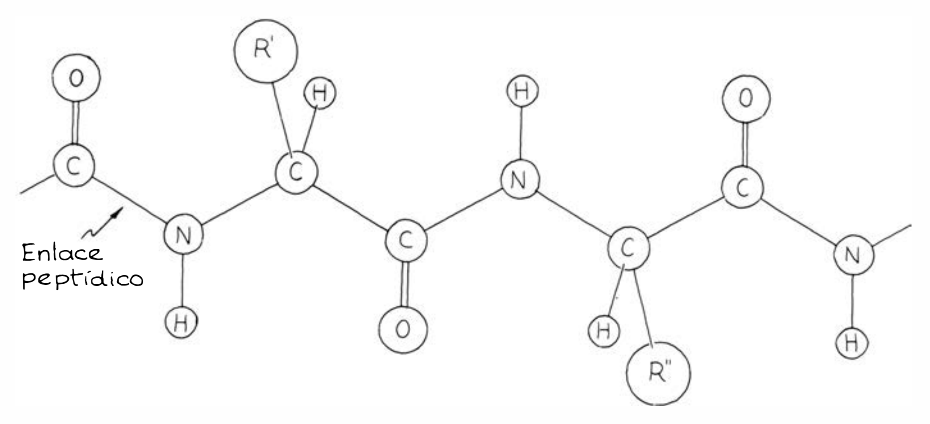
Cuando la mezcla se calienta, las burbujas de aire se dilatan y, por ello, dilatan el merengue. El calor desnaturaliza aun más las proteínas, favoreciendo su despliegue y permitiendo que la malla se ensanche a la vez que el aire se dilata. Asimismo, los enlaces entre las proteínas se multiplican y las claras se coagulan para formar una estructura firme. Lograda la dilatación, ya no son necesarias las burbujas de aire atrapado, porque el calor ha endurecido las claras al alcanzar éstas su estructura dilatada final. Si el merengue se calienta algo más es con la exclusiva finalidad de que, al secarlo, adquiera una consistencia quebradiza y crujiente.

El batido excesivo echa a perder el merengue porque desenreda demasiado las proteínas. A raíz de lo cual, la malla se torna muy sólida y el agua se separa de las proteínas. Señal de que la malla ha aumentado en exceso su firmeza la tenemos en la formación de crestas espesas al retirar la paleta de batir. En el caso de sobrada consistencia sólida, las claras no pueden dilatarse convenientemente durante la cocción al ritmo que lo hacen las burbujas de aire atrapado. Las burbujas revientan, y el merengue se viene abajo antes





*Dispersión de la luz láser por las claras de huevo antes (***izquierda***) y después (***derecha***) de calentadas*



Cadena de aminoácidos y enlaces peptídicos en las proteínas de clara de huevo

de que el calor haya tenido tiempo de coagular las proteínas; para cuando llegue la coagulación, el merengue ya se habrá desinflado y el cocinero se encontrará con claras de huevo cocidas sobre una base de limón.

Vigilando atentamente la superficie, un buen cocinero sabe cuándo se está al borde de batir las claras más de la cuenta, y acecha el momento crítico en que empiezan a perder espumosidad y están a punto de formarse gotitas de agua. La proteína desenredada es soluble en agua; en el exterior del grumito que forma existen numerosos lugares hidrófilos (que atraen agua), a los que pueden unirse moléculas de agua a través de enlaces hidrófilos. La mayoría de los lugares hidrófobos (que repelen el agua) están ocultos en el interior del grumito de proteína; entonces, cuando la proteína se desenreda, quedan al descubierto más lugares hidrófobos y escasean los hidrófilos para enlazar moléculas de agua.

Si el despliegue de las proteínas avanza todavía más, su enlace con el agua se ha reducido tanto que pierden su solubilidad. Entonces, si tenemos una cantidad pequeña de proteína dentro de agua, aquélla precipitará. Pero en el caso de un cuenco con claras de huevo, será el agua lo que se verá separarse y formar gotitas. Esto advierte al cocinero de que las proteínas se han desnaturalizado en demasía; revela, además, que el agua ya no es capaz, como antes, de mantener burbujitas de aire en la malla de gel, dado que las hidromoléculas no están ya unidas a las proteínas que forman dicha malla. Así pues, la separación de agua de las claras es señal de que el merengue no saldrá bien.

Cadenas de aminoácidos
Entorno hidrófobo (no atrae moléculas de agua)
CH_3
S—S
Enlace bisulfuroso
OH
Grupo polar (atrae moléculas de agua)
Negativo
Positivo
CO — NH_3
Enlace iónico
Negativo
COO H^+ Na^+
Entorno negativo (atrae iones positivos)
CH_3 CH_3 CH_3
Enlace hidrófobo o de van der Waals
Positivo
NH_3 OH^- Cl^-
Entorno positivo (atrae iones negativos)
OH—OH
Enlace hidrogénico

Tipos de enlaces en proteínas

Las claras de los huevos frescos son transparentes, pero se tornan opacas cuando sus proteínas se desnaturalizan. Para poner de manifiesto este hecho preparé una experiencia consistente en colocar un cubilete lleno de claras sobre un calientaplatos y traspasarlo con un haz luminoso estrecho procedente de un láser de helio-neon. Inicialmente, el haz se distinguía con toda nitidez dentro de las claras, en parte a causa de que las proteínas globulares dispersaban algo de luz. Aparecía una dispersión por no ser homogéneas las claras. Pero una vez que el calientaplatos adquirió cierta temperatura y facilitó la coagulación de las proteínas en malla, el haz perdió nitidez. Cuando finalizó la coagulación, las claras absorbían la mayor parte de la luz.

La investigación ha demostrado que pueden confeccionarse merengues a partir de mezclas que contengan únicamente una o dos de las proteínas que normalmente se encuentran en las claras de huevo. Sólo con ovalbúmina, el merengue tiene casi el mismo volumen, pero presenta una textura más basta; además, se necesita más tiempo para que el batido produzca la espumosidad correcta. Sólo con globulina y ovomucina, el batido no necesita tanto tiempo, pero el merengue se derrumba después de la cocción, debido seguramente a que la malla carece de firmeza suficiente para aguantar el peso. La globulina sola produce una espuma de burbujas pequeñas, dotando así al merengue de una textura más fina que la que se consigue con ovalbúmina sólo.

Desconozco por qué la ovalbúmina y la globulina originan espumas de tipos diferentes, pero me atrevería a afirmar que ello está relacionado con la facilidad con que puede desnaturalizarse cada proteína. Así, como la ovalbúmina tarda más que la globulina en producir una espuma de calidad, la primera opone quizá mayor resistencia a ser desenredada por batido; es casi seguro que esta característica se deba a una estructura terciaria que en la ovalbúmina sea algo más resistente que en la globulina. Como la globulina parece desplegarse con mayor facilidad, una mezcla de globulina produce proteína desenredada más deprisa que una mezcla de ovalbúmina y, con mayor cantidad de globulina desenredada, la malla de gel será más fina.

Cuando un cocinero bate claras de huevo naturales, la textura de la espuma se debe más a la globulina que a la

ovalbúmina. Aunque la ovomucina no espuma, parece que su malla favorece la estabilización de la espuma resultante de la desnaturalización de la globulina y la ovalbúmuna. Los cocineros saben que los huevos de pata dan unos merengues de escasa calidad; la razón de ello quizás estribe en que contengan menos globulina que los de gallina.

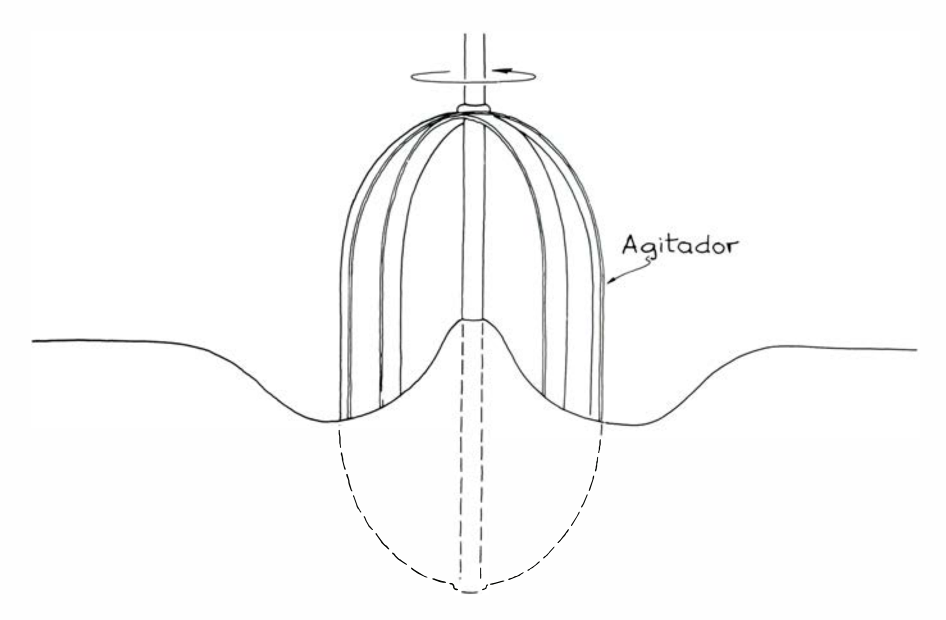
Antes de separar las claras de las yemas, los huevos deben encontrarse a la temperatura ambiente. La mayoría de los libros de cocina previenen contra la utilización de huevos tomados directamente del refrigerador, porque las claras extraídas de huevos fríos dan una espuma insignificante. El defecto reside en la viscosidad, dado que, cuando los huevos están fríos, su viscosidad es elevada y el batidor no puede entremezclar convenientemente burbujas de aire dentro del fluido.

Cuando accidentalmente se añade yema a las claras, las moléculas de lípidos (grasas) se fijan a los lugares hidrófilos mediante fuerzas de van der Waals. Por ello, tales lugares ya no están libres para enlazarse con otras proteínas y formar una malla de gel. Será entonces más difícil atrapar aire y menguará el volumen alcanzable batiendo las claras. A causa del aire atrapado, cuando bato claras de huevo adecuadamente, el volumen que ocupan aumenta cinco o seis veces. Pero si agrego un trocito de yema (u otro foco de grasa cualquiera) a la misma cantidad de clara, la dilatación se reduce.

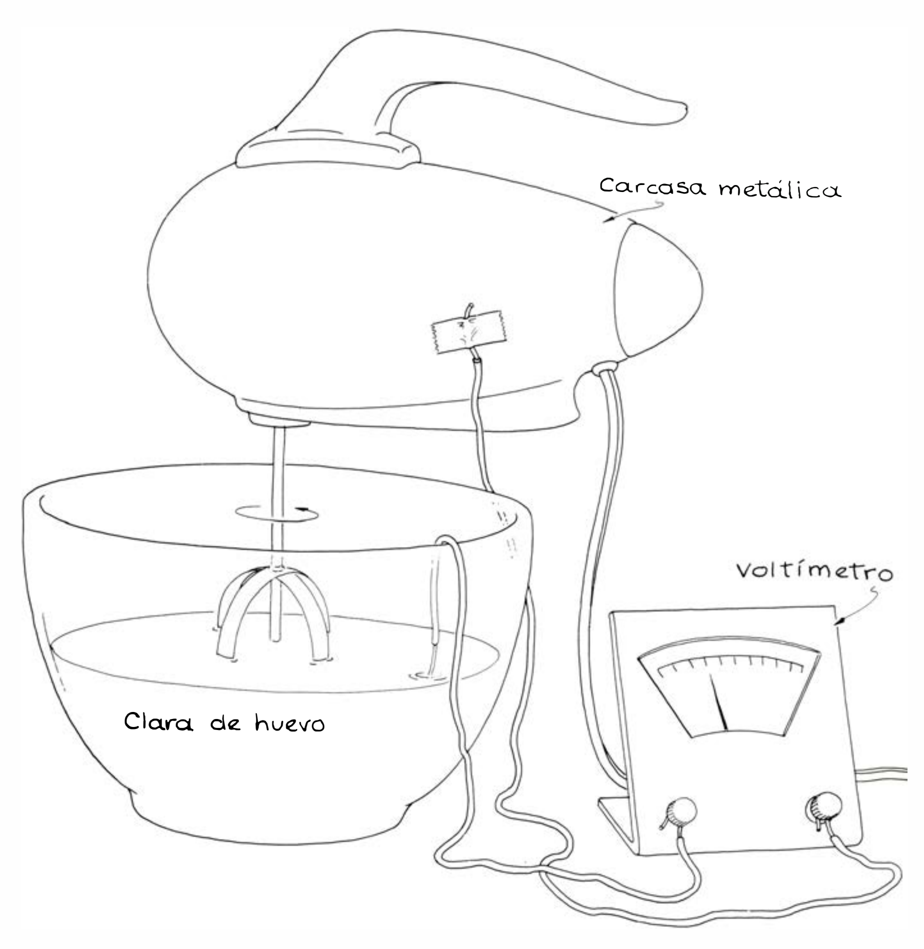
Para comprobar este efecto se baten las claras de tres huevos de primera calidad y, cuando la superficie reluzca convenientemente, se hace una señal en la pared del cuenco que indique la altura de las claras. Se enjuaga y se seca el cuenco; se separan las claras de otros tres huevos. A éstas se añade ahora el 10 por ciento aproximadamente de una de las yemas, y se bate como antes. Se observará que, tras batir prolongadamente, el volumen viene a ser la mitad que el de la tanda precedente y que la mezcla es notablemente coladiza, lo que revela la falta de una buena malla de gel.

He advertido otro indicio de ausencia de malla, o de su calidad deficiente: al batir claras sin trazas de yema, la mezcla trepa por el eje de mi batidora eléctrica. Se trata del efecto Weissenberg e indica que el fluido que se agita es no newtoniano. (En el número de enero de 1979 me ocupé ya de dicho efecto.) Lo que ocurre es que las mallas bien conformadas se agarran al eje, al girar éste, y se elevan por él; pero si las claras contienen yemas, la malla es deficiente y apenas trepa por el eje.

Hay cocineros que creen que añadiendo un ácido débil se facilita la labor de formar espuma al batir la clara. El ácido se añade directamente en los primeros instantes del mezclado, o bien se pone en el interior del cuenco antes de introducir las claras. Quizá sea válida esa apreciación. Al agregar el ácido, sus moléculas se disocian en iones positivos y negativos. Los iones positivos tienden a reunirse en torno a los lugares de las proteínas con carga negativa. Supongamos ahora que estos lugares intervienen en el mantenimiento de la



Efecto Weissenberg en el batido de claras de huevo



Montaje para comprobar si se desarrolla una corriente eléctrica al batir claras de huevo

Un molde de 20 cm con masa cocida

Merengue:
4 claras de huevo (1/2 copa) a temperatura ambiente
1/4 cucharadita de crémor
1/2 copa de azúcar

Base:
1/3 de copa de almidón de maíz
1 1/2 copa de azúcar, superfina
1/4 de cucharadita de sal
4 yemas de huevo, levemente batidas
1/4 de copa de zumo de limón
2 cucharadas de piel de limón rayada

1 cucharadita = 5 mililitros
1 cucharada = 15 mililitros
1 copa = 237 mililitros

Ingredientes para una tarta de merengue al limón

estructura enmarañada de la proteína; entonces, con los iones positivos agolpados en su entorno, no pueden desempeñar esa función en la estructura y la proteína se desenreda. Análogamente, los iones negativos pueden arracimarse en torno a los puntos de la proteína cargados positivamente y favorecen asimismo el despliegue de éstas.

La adición de un ácido posee un efecto aun más importante, crítico para la formación de la malla de gel. Las proteínas de las claras recién preparadas tienen una carga neta negativa, por cuya causa forman una malla con cierta dificultad, habida cuenta de su mutua repulsión. Cuando se añade un ácido, se reduce la carga neta de cada proteína en virtud del agolpamiento de iones en torno a los lugares cargados. Mengua la repulsión entre las proteínas, pudiéndose enlazarse entre sí con mayor facilidad para formar malla.

Si se agrega sal (NaCl) a las claras antes de batirlas, pueden suceder dos cosas. La sal se disocia en iones positivos (Na^+) y negativos (Cl^-), que se agolpan en torno a los lugares cargados de las proteínas. Los iones de sodio se agolpan en los puntos negativos y los de cloro en los positivos. Si tales lugares tienen algo que ver con la estructura terciaria de las proteínas, éstas comenzarán a desenredarse. Este efecto es de alcance reducido, ya que la mayoría de las proteínas no se desnaturalizan con la sal.

El agolpamiento de iones junto a los lugares cargados también puede reducir la carga neta de las proteínas, con lo que disminuye su repulsión mutua y es más fácil que se forme una malla.

Pese a estas posibles ventajas derivadas de la adición de sal, el efecto más importante de ésta es indeseable, ya que puede eliminar agua de las proteínas. Como resultado, el agua ya no podrá formar película alrededor del aire introducido en la mezcla por el batido y no se fijará a la malla de proteínas entrelazadas que se trata de crear. Entonces, al no retener aire, habrá muy poca espuma. Si una mezcla de claras de huevo necesita sal, ésta sólo debe añadirse después de batir y obtener la espuma. Incluso entonces, todo lo que sea añadir algo más de una cantidad muy pequeña hará que la espuma tienda a desinflarse.

¿Y qué decir acerca del cobre del cuenco? Quizás intervenga de algún modo, pero no acierto a ver cómo. El aluminio se comporta deficientemente porque decolora las claras. Por su parte, el cobre podría favorecer el proceso si disminuyese la carga neta de las proteínas, facilitando así la formación de la malla, pero no lo creo. Alguien podría acometer el experimento y observar si los cuencos de cobre hacen más fácil el batido de las claras de huevo, aumentan el volumen de las claras batidas correctamente o aumentan el volumen del merengue.

En muchos libros de cocina se lee que las claras de huevo se espesan, en parte, a causa de un campo eléctrico que se genera con el batido. Los recetarios ofrecen instrucciones acerca de los mejores procedimientos para crear dicho campo, o no destruirlo por lo menos. En uno que llegó a mis manos se decía que debía evitarse batir las claras en un cuenco húmedo para que la película de agua no destruya este campo.

El efecto cortante creado por un batidor o una paleta al moverse podría originar un campo eléctrico muy similar al que se forma cuando se tira del extremo de un rollo de celofán. Entonces, si realmente se genera un campo eléctrico en el seno de las claras de huevo, éste podría favorecer el desenredo de las proteínas rompiendo algunos de los enlaces más débiles de la estructura terciaria.

Cuando yo pretendí detectar un campo eléctrico en claras de huevo batidas en mi aparato eléctrico, fijé uno de los bornes de un voltímetro a la envuelta de la batidora y uní el otro a un trozo de cable conductor que introduje en las claras. Las sondas registrarían así todo campo eléctrico que pudiera hallarse entre los agitadores y el cable inmerso en las claras. Pero el instrumento no registró tensión alguna, al menos para su ajuste de escala más sensible, que es de 0,6 volt. (Si alguien intenta esta experiencia, que se asegure de que el cable no se enreda en los agitadores durante el giro.)

Fijé luego un cable a la primera sonda y lo hice descender dentro de las claras y cerca del otro cable. De nuevo perdí el tiempo esperando una diferencia de potencial entre ambos cables. Así pues, si realmente la mezcla genera alguna tensión, ésta debe ser inferior a 0,05 volt y estará ubicada precisamente junto a las claras que en cada instante se están hendiendo. Puedo afirmar, pues, que en la masa de claras no se forma campo eléctrico. De lo que se deduce que no hay cocinero que pueda poseer dominio alguno sobre el campo eléctrico (si existe) y, por tanto, no debe preocuparle en absoluto.

Se agrega el azúcar tras haber batido convenientemente las claras para obtener una espuma estable. Si se echa antes de formarse la malla de gel, se desprenderá el agua de las proteínas y el fluido se espesará, lo que haría más difícil airear la mezcla. Los buenos cocineros no añaden, pues, el azúcar hasta que el batido haya formado la espuma, y aún entonces la vierten lentamente mientras siguen batiendo. No sirve el azúcar de granulado ordinario; se necesita azúcar superfino que se disuelva rápidamente en las paredes acuosas de las burbujas de aire. El azúcar de granulado ordinario tarda demasiado en disolverse y forma terrones en el merengue. Para observar el progreso de la disolución se prueba la mezcla. Una textura áspera delata azúcar no disuelto. Si es necesario, puede añadirse una pequeña cantidad de agua para completar la disolución.

Cuando en el merengue queda azúcar no disuelto, en su superficie apare-

cen gotitas de jarabe al cocerlo. Estas tartas se llaman merengues llorones. (Quizá por las lágrimas del cocinero.) Cocida la tarta, ésta debe encerrarse en un recipiente, pues, de no hacerlo, el azúcar absorbería agua del aire provocando la formación de gotitas de agua en la superficie de la tarta. Estas gotitas no sólo añaden un peso que puede hacer que el merengue se venga abajo, sino también ablandar la malla de gel proteínica que constituye su armazón. Por esta causa, hube de renunciar en cierta ocasión a la confección de una tarta de merengue al limón durante el húmedo verano de Cleveland, cuando el aire contiene tanto vapor de agua que el merengue comienza a derrumbarse casi al instante de sacarlo del horno.

La base de limón para una tarta de merengue al limón se prepara con yemas de huevo, azúcar, almidón de maíz, sal, mantequilla, zumo de limón y piel de limón rayada. Para prepararla se empieza calentando el almidón, el azúcar y la sal dentro de una cacerola, a la que se añade agua mientras se agita. Cuando esta mezcla está suave y caliente, se retira la cacerola del fogón y se vierte la mitad del líquido en un cuenco de mezclar que contiene las yemas, que ya se habrán mezclado ligeramente. Tras mezclarlo ahora todo rápidamente, la masa se devuelve a la cacerola. Entonces, se agita y se calienta toda la mezcla a calor mediano hasta que haya hervido alrededor de un minuto. Seguidamente, se añaden el zumo, las pieles y la mantequilla, a la vez que se agita, y la mezcla se vierte sobre un soporte de tarta ya cocido.

Así reza el procedimiento clásico para confeccionar rápidamente la base. Los buenos cocineros ponen más cuidado, sobre todo en la textura de la tarta. Se me permitirá simplificar concentrándome en las yemas y en el azúcar. El almidón espesa la base y las pieles dan a la tarta su consabido sabor; pero el secreto de la base reside en las yemas y en el azúcar.

Las yemas y el azúcar forman un jarabe cuando se baten juntos. El batido aligera la mezcla desnaturalizando las proteínas de las yemas y, a pesar de que las grasas que éstas contienen obstruyen la formación de una malla de gel a base de dichas proteínas, sí se forma cierta malla. Este batido introduce además burbujas de aire en la mezcla. El azúcar absorbe agua de las yemas y forma jarabe, por cuya razón el producto adquiere espesor, si bien se aligera un tanto merced al aire que el batido introduce entre las proteínas desnaturalizadas. Cualquier experto puede averiguar si la mezcla se ha batido correctamente examinando cómo se escurre por un batidor de alambre cuando éste se saca de la mezcla. Si la mezcla se repliega sobre sí misma al retornar al cuenco, es que está lista para verter sobre el soporte de la tarta.

Los libros de cocina abundan en instrucciones sobre la cocción de la tarta. Los hay que recomiendan una cocción corta, de siete a nueve minutos, a la temperatura, relativamente elevada, de unos 200 grados Celsius. Otros piden una temperatura mucho más baja (80 grados Celsius) y un tiempo de cocción de varias horas, tras de lo cual se apaga el horno y la tarta se deja en él toda la noche. Y, por fin, están los que aconsejan temperaturas y tiempos de cocción intermedios.

En general, una temperatura baja produce un merengue terso y crujiente, y una temperatura alta lo hace gomoso y masticable. En mi opinión, las descripciones de los libros pecan de imprecisión. Sería más exacto referirse a la elasticidad del merengue, que viene determinada por la temperatura. A baja temperatura y con un tiempo de cocción largo, el merengue se infla lentamente al dilatarse las burbujas de aire y se seca. Se evita así que adquiera excesiva elasticidad, en parte por el secado, pero sobre todo debido a que la malla de gel se estira durante tanto tiempo que se vuelve rígida. A mayor temperatura y con una cocción más breve, el merengue se estira rápidamente y quizá no se seque tan a fondo; en cuyo caso, será más elástico porque la malla no sufre tracción durante tiempo suficiente para volverse rígida.

En todas las recetas se exige que el calor proceda de la base del horno, nunca de lo alto, al objeto de impedir que la parte superior del merengue absorba directamente la radiación infrarroja y se torne marrón (o negra) en seguida. Para algunos cocineros, la tarta debe cocerse hasta que su parte más alta comience a ponerse marrón; otros afirman que debe evitarse este color.

Me agradaría mucho contarles más cosas acerca de los merengues, especialmente de las fuerzas principales responsables de la malla de gel proteínica y del modo en que las afectan los distintos ingredientes. Tengo en proyecto continuar el estudio de otros platos de huevos en un artículo futuro. En particular, me dedicaré a los secretos ocultos en las magníficas tortillas, suflés y huevos revueltos de mi abuela (insuperable con los merengues y con las tartas de merengue al limón), elaboraciones todas ellas imposibles para mis manos chapuceras.

Libros

Tres líneas de novísima investigación: teoría de sistemas, física estadística y bioingeniería

F. Sáez Vacas, Manuel García Velarde y Miguel Aguilar

TEORÍA GENERAL DE SISTEMAS. (UN ENFOQUE METODOLÓGICO.), por George J. Klir. Versión española de F. J. Valero López, con introducción y revisión técnica de Eduardo Bueno Campos; Ediciones ICE; Madrid, 1980; 383 páginas. Dentro del movimiento actual en la investigación de sistemas generales, G. J. Klir ocupa un lugar relevante. Hasta hace poco tiempo director ejecutivo de la Society for General Systems Research, hoy es su presidente electo, además de presidente del Comité Ejecutivo de la International Federation for Systems Research. Sus actividades editoriales abarcan, entre otras, las de editor del *International Journal of General Systems* y las de editor-jefe de una colección de libros bajo el título genérico de Frontiers in Systems Research. Es profesor y dirige el departamento de ciencia de sistemas de la Universidad estatal de Nueva York en Binghamton. Ha escrito, coescrito o editado numerosos artículos y varios libros sobre circuitos, modelos cibernéticos y sistemas generales. (En una bibliografía sobre investigación básica y aplicada de sistemas generales realizada en 1977, y que comprende publicaciones de 1084 autores diferentes, figura Klir en segundo lugar, en cuanto al número de trabajos publicados, de una lista que se ordena así: Ashby, Klir, Rosen, Mesarovic, Zadeh, Bertalanffy, Arbib, Rapoport, Simon, Churchman, etcétera.) El libro que nos ocupa data de 1969 y es un clásico de este movimiento investigador.

Puesto que el libro de Klir describe una teoría de sistemas generales (más exactamente un "Approach to General Systems Theory", como indica el título original) y una metodología, cualquier reseña del mismo debería abarcar ambas vertientes, a saber, la teoría como tal y el soporte en que se nos sirve, es decir, el propio libro. Desde el punto de vista del sujeto lector se produce siempre una evidente interrelación de esas dos facetas, en donde la presentación que de la teoría se haga condicionará fuertemente su grado de asimilación y aceptación. (Es obligado introducir esa coautoría delegada que corresponde al trabajo de traducción.) Considerando ambos factores, estamos ante un libro importante, un libro difícil, un libro inacabado y un libro hasta cierto punto mal construido.

La evidencia de que se trata de un importante y clásico libro sobre la materia se obtiene por mera constatación frecuencial de su referenciación en la literatura científica de este área. Las causas de tal éxito radican en la enjundia intrínseca o riqueza de la teoría compuesta por Klir, a través de una sabia disposición y conexión de materiales propios y ajenos. A que su éxito haya podido ser, tal vez, menor del merecido puede que no sean ajenas las calidades formales del libro.

Por limitaciones de espacio, no es factible detenerse en glosar cada uno de los puntos de la teoría de Klir que merecían atención. Cabe destacar su más esencial y original valor: la jerarquización de los sistemas en niveles epistemológicos. Frente a la concepción estructural generalizada de los sistemas (el sistema como ensamblaje integrado de elementos interactivos o cualquiera de las definiciones derivadas), Klir parte de la realidad polifacética de la tarea de aproximación al conocimiento y de las trabas con que se encuentra, empezando por las que representan el mismo observador-investigador y la complejidad del objeto motivo de investigación. Según esto, la definición estructural típica es abusiva la mayoría de las veces, por lo que Klir propone, no una, sino cinco definiciones básicas.

La primera y primaria, por cuanto que supone el menor nivel de conocimiento, es aquella que se inspira en una vieja definición de Ashby (*Design for a Brain*, 1952), que decía: "se define un sistema como cualquier conjunto de variables que el observador selecciona de entre las disponibles en la 'máquina' real". Klir traduce: "sistema es el conjunto de las cantidades externas más el nivel de resolución". El observador delimita un objeto (un avión, un grupo social, una ciudad, ..., la 'máquina' real); sobre él selecciona unas cantidades, que puede medir u observar externamente, de acuerdo con cierto poder de resolución instrumental, y tiene un sistema. Si cambia alguna de esas cantidades o el nivel de resolución de alguna cantidad, está observando otro sistema. Y el conocimiento transita obligatoriamente por esta vía.

Un segundo nivel de sistema que recubre al anterior, y al tiempo está esclavizado a él, es el de sistema visto como actividad. Actividad es "el conjunto de las variaciones temporales (durante un intervalo) de las cantidades consideradas en el nivel de resolución dado". Por esta definición vemos al objeto manifestándose a lo largo del tiempo, pero su manifestación se observa a través de la óptica deformadora del 'sistema' (definición de nivel 1).

El resto de definiciones, 'comportamiento permanente', 'estructura real de universo y acoplamientos' y 'estructura de estados y transiciones' guardan entre sí y con las dos primeras una relación de jerarquía, de un lado en esa dependencia señalada en cuanto al conocimiento obtenible en la aproximación al objeto y, de otro, en lo que se refiere a la configuración pragmática de ese acercamiento cognoscitivo. Y con esto último, nos referimos a los aspectos metodológicos de la teoría de Klir, entendiendo aquí el término 'metodología' en un sentido amplio.

Klir resume y propone un conjunto de conceptos y procedimientos para resolver tres grandes y arquetípicas clases de problemas: el problema del análisis de un sistema, el problema de la síntesis de un sistema y el problema de la caja negra, por este orden de dificultad. El primer arqueproblema ha sido históricamente más trabajado y no es casualidad que su método general de resolución tome como punto de partida

la definición estructural típica de sistema, como tampoco lo es que los otros dos problemas se inicien, respectivamente, en la definición de nivel 2 (actividad) y en la definición de nivel 1. Así pues, pivotando de manera ostensible sobre la subjetividad y relatividad epistemológica de esta primera definición, construye tres edificios metodológicos (los tres problemas puros) que consisten en los procesos de adquisición de informaciones (con especificación clara de sus límites y carencias) para pasar de una a otra definición de 'sistema'. Siendo uno el objeto y muchos los problemas con él planteables en principio, no es fácil resaltar de forma más nítida la necesidad de contar con varias definiciones distintas de 'sistema' y con sus canales de intercomunicación (operativamente desarrollados en la teoría de Klir casi en exclusiva para sistemas discretos, dicho sea de paso).

Este es un libro difícil, e incluso muy difícil, si se acepta que bajo su apariencia oculta tesoros que requieren esfuerzos de reflexión para llegar a ellos y para desarrollarlos. Además del nivel metodológico de la teoría, que no se acomoda bien a las necesidades de unos y otros, su dificultad emana del grado de inmadurez de la propia teoría y de la imperfección didáctica con que ésta se describe. Diríamos que los científicos sociales han sido preferentemente atraídos por esta teoría; ello tiene que deberse, sin duda, a sus atractivos conceptuales y a sus calidades epistemológicas, ya que desde un punto de vista metodológico su capacidad de descripción y resolución de problemas sociales es más que discutible (en el estado de formulación que presenta el libro). Por añadidura, su nivel de formalización resulta fuerte –y esto lo hemos contrastado con diferentes estudiosos sociales– para el que éstos habitualmente manejan, lo que unido a que la mayoría de ejercicios, ejemplos y fundamentos del libro proceden de la cibernética, de la electrónica y de la teoría de circuitos, plantea problemas indudables de adaptación real a las disciplinas sociales.

De otro lado, si se miran las cosas desde la óptica de los estudiosos de la electrónica, de la informática y de la automática, parecería como si, en su conjunto, éstos hubieran sido mucho menos o muy poco seducidos por la teoría general de sistemas y por la teoría de Klir, en particular. Son pocos quienes aprecien las excelencias y las posibilidades que esta teoría ofrece en cuanto a la ampliación del campo conceptual: visión de conjunto, límites de los métodos concretos, transdisciplinariedad..., y muchos más aquellos que la califiquen de superficial e inútil.

Cuando arriba calificaba a este libro de libro inacabado lo hacía metafóricamente, puesto que me refería a la teoría que contiene. Al lector de esta revista y al lector del libro les interesará saber que la elaboración de Klir no concluyó en el año 1969 con la publicación de su obra. Tanto él como otros autores han continuado desarrollando conceptos y métodos, de los que cabe citar la sexta definición de sistema (véase Orchard, en *Trends in General Systems*, editado por G. J. Klir, Wiley N. Y. 1972; vertido al castellano con el título de *Tendencias en la teoría general de sistemas,* Alianza Universidad, 1978), el programa de investigación en metodología de sistemas del GSPS (General System Problem Solver) en la Universidad de Binghamton y una aplicación más decidida orientada a las ciencias sociales (véase R. E. Cavallo en *The Role of Systems Methodology in Social Science Research*, Martinus Nijhoff Publishing, Boston 1979).

El nivel pedagógico de un libro no se mide por el número de ejemplos, sino por la adecuación de éstos a lo que se quiere comunicar (entre otros factores), y en esto el libro de Klir deja bastante que desear. Su primer y fundamental capítulo, que es donde desarrolla prácticamente toda su teoría, es denso en exceso, de digestión muy pesada y diríase que casi impenetrable para un lector no iniciado, salvo que vuelva a él una y otra vez después de estudiar los capítulos sucesivos. Los ejemplos empleados, con ser numerosos, generalmente resultan triviales (y por tanto poco convincentes), en relación con el supuesto y ambicioso empeño de la teoría general de sistemas. El nivel de formalización desbordará a algunos lectores y a quienes tienen mayor formación matemática les parecerá corto, aunque en ningún momento esta teoría se reclame de teoría matemática de sistemas generales. Al libro le faltan dosis de "marketing", sobre todo en lo que se refiere a glosar más pormenorizadamente diferencias y similitudes con otros enfoques y ventajas con respecto a teorías más estrechamente disciplinadas. Le sobran unas cuantas erratas, no pocas de ellas escondidas en algunos ejemplos, lo que, por consiguiente, no facilita su estudio. Hay que decir que el traductor español las ha respetado escrupulosamente casi todas.

Ello nos lleva a dedicar unos párrafos a su publicación en lengua castellana. Sea muy bienvenida, porque éste es el primer libro básico sobre teoría y meto-

dología de sistemas generales publicado en español. Cabe, no obstante, hacerle unas breves puntualizaciones. La primera es que esta obra es anterior en su cronología a *Tendencias en la teoría general de sistemas*, siendo esta última complementaria de aquella. La segunda es sobre el título que, además de ser incorrecto o incompleto en la portada: *Teoría general de sistemas*, tiende a alimentar una inflación y una confusión respecto al fondo bibliográfico sobre este tema. Ya contamos con otros dos libros, muy diferentes en el contenido y en la forma, titulados casi idénticamente: *Teoría general de los sistemas* de L. V. Bertalanffy, Fondo de Cultura Económica, México, 1976, y *La teoría general de sistemas* de P. Voltes, Editorial Hispano Europea, Barcelona, 1978.

Le sugerimos al editor que añada una fe de erratas, pues el traductor, aunque en puridad no tiene por qué enmendarle la plana al autor y corregir los defectos del original, deja traslucir nítidamente que su profesión no tiene nada que ver con la ingeniería eléctrica o electrónica. Tal vez por eso se ha mostrado insensible a erratas en ejemplos como el 1.7 (fig. 1.3), el 1.13, el 3.6, el 3.15 (fig. 3.21), el 4.2, el 6.4., ..., llama relais al relé, gatillo al disparador (trigger), compuertas a las puertas, elementos de demora a los retardos (delays), sistemas de pulso a los sistemas pulsados... Sin embargo, ciertos detalles, fáciles de cuidar, habrían merecido el agradecimiento de los lectores, como utilizar las letras r, a, v y no r, y, g, por rojo, amarillo, verde, en el ejemplo 1.11, o dirigir bien la punta de flecha del diagrama de la figura 6.11, etc... Algo que llama la atención es que se traduzca sistemáticamente 'environment' por 'ambiente' y no por 'entorno', como parece más corriente cuando se aplica genéricamente a toda clase de sistemas. Por último, se ha eliminado el índice alfabético del original, ahorro mínimo en un plano económico que gravitará pesadamente sobre cada lector.

En resumen: teoría importante iniciadora de una línea de investigación sobre metodología de sistemas generales hoy en pleno estado de desarrollo, servida en un libro no demasiado conseguido; libro inevitable, no obstante sus defectos propios y los añadidos, en la biblioteca de toda persona interesada en el movimiento sistémico, cualquiera que sea su campo disciplinario de procedencia. Actualmente se utiliza como texto en la Escuela Técnica Superior de Ingenieros de Telecomunicación y en la Facultad de Informática, ambas de la Universidad Politécnica de Madrid. (F. S.)

FIELD THEORY, QUANTIZATION AND STATISTICAL PHYSICS (a la memoria de Bernard Jouvet), editado por E. Tirapegui; publicado por Reidel, Dordrecht (Holanda); 348 páginas. Se recogen aquí diecisiete trabajos dedicados a la memoria de Bernard Jouvet, físico francés nacido el 5 de febrero de 1927 y muerto el 8 de enero de 1978. La obra consta de dos bloques: teoría de campos (en sentido amplio) con 12 trabajos, que cubren 216 páginas, y física estadística con 5 trabajos, que se extienden a lo largo de 91 páginas. Esta distribución no corresponde a la que, de hecho, presenta el libro; pues en la primera parte se incluye (haciendo 14 en su total parcial) uno sobre inestabili-

dades en plasmas, en el que se estudia el papel de la viscosidad y otros efectos disipativos en un problema de magnetohidrodinámica. El volumen posee el atractivo, sin embargo, de puentear solapando los temas de los dos bloques, como corresponde a la "unificación de facto" existente actualmente en física, allende nuestras fronteras.

En la parte I se abordan cuestiones de teorías renormalizables con ruptura espontánea de simetría (tal y como ocurre en fenómenos críticos y transiciones de fase en física estadística), cálculos (algebraicos y numéricos) en electrodinámica cuántica (QED), gran unificación y gran explosión, gravitación y relatividad general, partículas autocompuestas (un término, al parecer, inventado por Jouvet) etcétera. En la parte II, con el artículo antecitado sobre plasmas incluido, se discuten cuestiones de inestabilidades hidrodinámicas y turbulencia (geometría cilíndrica en magnetohidrodinámica, convección de Bénard-Rayleigh y otros), aparición de "caos" (y "estocasticidad") en sistemas dinámicos deterministas sencillos (desde la perspectiva tanto de la teoría de las transiciones de fase como del moderno-actual enfoque de la turbulencia en física de fluidos), efectos de fluctuaciones en ecuaciones macroscópicas (Navier-Stokes, Langevin, etcétera) y transiciones de fase en modelos de espines sobre retículo.

Casi todos los trabajos aquí compilados me parecen legibles para un estudiante avanzado de doctorado y recomendables para los aspirantes a profesar la física teórica. Algunos aportan resultados originales o hacen un análisis de la literatura relativa al problema tratado. Hay que señalar, sin embargo, que mucho del material es "formalista" y, el todo, bastante heterogéneo, como corresponde a un libro donde lo que primó, pienso, fue la relación personal con Jouvet. La heterogeneidad es buena si se toma como medio de "alfabetizarse" o como apertura a la "versatilidad", o fluencia en la "lengua".

Se incluye una calurosa biografía de Jouvet, su trayectoria docente o investigadora, lista de publicaciones (40 relativas a la parte I del libro y 4 sobre transiciones de fase, parte II) y un recuerdo emocionado a su viuda (recuerdo escrito en francés cuando el libro entero está en inglés).

Cabe decir que tanto la física estadística como la física de fluidos son áreas de investigación y docencia, en España, que junto con la física atómica y molecular (y no nuclear) tienen un desarrollo oficial especialmente deprimido, si lo comparamos con su relevancia en otras longitudes y latitudes (hacia el Norte) en los tiempos que corren. Oficialmente, sólo hay un catedrático (numerario) de mecánica estadística. (El nombre de física estadística es legalmente, si no imposible, sí al menos difícil de introducir en nuestro sistema docente universitario.) No existe ninguna cátedra de física de fluidos en universidades no politécnicas (o sea, en ninguna sección de físicas, si excluimos la específica física del aire, directamente ligada a la meteorología), y departamento de física de fluidos sólo ha habido uno (y por tiempo breve) en todo el sistema universitario no politécnico. Afirmar que la física estadística es una de las más completas, importantes y útiles subdisciplinas de la física quizá sea decir algo "acertado en una socie-

dad injusta" (arrimo el ascua a mi sardina y me apoyo en Quevedo) y, por tanto, peligroso.

Por último, y volviendo al libro objeto de esta recensión, el precio de este volumen se eleva a cuarenta y siete dólares. Me parece un precio fuera del alcance de nuestros estudiantes de doctorado o de nuestros investigadores o profesores. Cabe aconsejarlo para bibliotecas de facultades de ciencias o de físicas y matemáticas, tanto por sus posibilidades "educadoras" y "alfabetizadoras" como por su valor intrínseco de investigación y, marginalmente, por su esmerada presentación y agradable tipografía. (M. G. V.)

INSTRUMENTACIÓN Y MEDIDAS BIOMÉDICAS, por J. Cromwell, F. S. Weibell, E. A. Pfeiffer y L. B. Usselman; XIII; 427 páginas. Barcelona; Boixareu Editores, 1980. (Traducción de R. Pallás de la obra del mismo título publicada en inglés en 1973.) A pesar de su corta vida –la primera edición americana es de 1973– puede decirse que se trata de una obra ya clásica en el campo de la bioingeniería. La bioingeniería, como los autores señalan al comienzo del libro, no hay que olvidarlo, es un campo interdisciplinar de desarrollo muy reciente. Ambas facetas del mismo, su carácter interdisciplinar y su juventud hacen que no existan apenas libros que se ocupen de ella, considerándola dotada de personalidad propia y abarcando todos sus aspectos y características.

Es ese vacío de textos escritos sobre el tema de la bioingeniería como conjunto el que esta obra pretendió llenar, en su día, en el mundo de habla inglesa y pretende ahora con su versión castellana realizar en el nuestro. La primera de estas pretensiones se ha realizado plenamente, convirtiéndose en un clásico en la materia, rejuvenecido ya por una nueva edición, en lengua inglesa, que aparece en 1980 sólo con ligeras variantes estructurales y las puestas al día inevitables en un campo de tan rápida evolución. La segunda es de esperar que la cumpla igualmente ya que en nuestro país –y probablemente en los demás de habla española– se carecía de una obra como ésta, y hay que decir que está bien dotada para ello.

Consta de dieciséis capítulos seguidos de diversos apéndices. Los dos primeros abordan temas generales. En el primero, se definen, por un lado, los conceptos de bioingeniería, bioinstrumentación, biometría, etcétera, al tiempo que se establece y justifica la estructura de la obra. El segundo se ocupa de las relaciones entre el hombre y los instrumentos de medida desde el punto de vista de la ingeniería de sistemas, estableciendo el sistema hombre-instrumento.

Los catorce capítulos restantes se dedican, por este orden, al estudio de los potenciales bioeléctricos, electrodos, el sistema cardiovascular, medidas cardiovasculares, cuidado y monitorización de pacientes, medidas en el sistema respiratorio, medidas de variables físicas (temperatura, fuerza, desplazamiento, velocidad y aceleración), el sistema nervioso, medidas sensoriales y estudio del comportamiento, biotelemetría, instrumentación para el laboratorio químico, instrumentación para rayos X y radioisótopos, el computador en instrumentación biomédica y seguridad eléctrica del equipo médico.

El libro termina con una bibliografía, no muy abundante pero seleccionada, seguida de cuatro apéndices, de los cuales los dos primeros están dedicados a la terminología y glosario médicos y a un resumen de las principales medidas fisiológicas; el tercero presenta una serie de datos del funcionamiento pulmonar, mientras que el cuarto recoge una serie de ejercicios y problemas ordenados por capítulos. La obra se cierra con un índice onomástico muy completo.

Dirigidos, según confiesan los autores, a lectores con formación de ingenieros y escasos conocimientos de fisiología, los temas tratados se describen a nivel de exposición general, remitiendo a la bibliografía oportuna para una mayor profundización.

Por otra parte, como nunca hay rosas sin espinas, es preciso realizar algunas puntualizaciones de signo menos positivo sobre la obra castellana. La primera y principal se basa en el hecho de publicar en 1980 un libro de 1973 sobre un tema en rápido desarrollo. Así se echan de menos en el mismo técnicas que hoy son de uso cotidiano en cualquier gran hospital, e incluso en centros de menor entidad. Nos referimos principalmente a las técnicas de exploración por ultrasonidos y a la tomografía axial computarizada entre otras. En otro orden de cosas se echa a faltar la presencia en el libro del microprocesador, especie de factótum cuyo impacto en la instrumentación biomédica, al igual que en otras ramas de la ingeniería, no es preciso resaltar aquí. Hemos de decir también que la bibliografía resulta bastante anticuada y de difícil acceso. Finalmente, la presentación del libro es discreta, si bien su sintaxis es susceptible de mejora. Un libro valioso sin duda, para cuya traducción, quizá, se debiera haber esperado a la aparición de la segunda edición americana. (M. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

R. A. Morse y P. K. Visscher en *Psyche (Journal of Entomology)*, vol. 86, págs. 103-113; 1979.

REACTORES NUCLEARES REFRIGERADOS POR GAS

NUCLEAR ENERGY-THE NEXT PHASE. Peter Fortescue. Forum científico internacional sobre geopolítica y energía, Universidad de Miami, Centro de Estudios Teóricos, 1980.

OPERATIONAL AND MAINTENANCE EXPOSURE TRENDS FOR NUCLEAR POWER: PROCEEDINGS OF THE 1980 UTILITY/USER CONFERENCE ON THE HTGR. John L. Helm. Gas-Cooled Reactor Associates, 1980.

LA DESINTEGRACION DEL PROTON

BARYON AND LEPTON NONCONSERVING PROCESSES. Steven Weinberg en *Physical Review Letters*. vol. 43, n.° 21, págs. 1566-1570; 1979.

FROM ATOMS TO QUARKS: AN INTRODUCTION TO THE STRANGE WORLD OF PARTICLE PHYSICS. J. S. Trefil. Charles Scribner's Sons, 1980.

VARIETIES OF BARYON AND LEPTON NON CONSERVATION. Steven Weinberg en *Physical Review D*, vol. 22, n.° 7, págs. 1694-1700; 1980.

CONCEPTUAL FOUNDATIONS OF THE UNIFIELD THEORY OF WEAK AND ELECTROMAGNETIC INTERACTIONS. Steven Weinberg en *Science*, vol. 210, n.° 4474, págs. 1212-1218; 12 de diciembre de 1980.

ANTIBIOTICOS BETA-LACTAMICOS

ANTIBIOTICS. H. W. Florey, E. Chair, H. G. Heatley, M. A. Jennings, A. G. Sanders, E. P. Abraham y M. E. Florey. Oxford University Press, 1949.

THE BIRTH OF PENICILLIN. Ronald Hare. George Allen y Unwin, Ltd., 1970.

THE CHEMISTRY OF PENICILLIN. Dirigido por Hans T. Clarke, John R. Johnson y Sir Robert Robinson. Princeton University Press, 1970.

CEPHALOSPORINS AND PENICILLINS: CHEMISTRY AND BIOLOGY. Dirigido por Edwin H. Flynn. Academic Press, 1972.

BETA-LACTAMASES. Dirigido por J. M. T. Hamilton-Miller y J. T. Smith. Academic Press, 1979.

ARCHIBACTERIAS

ARCHAEBACTERIA. C. R. Woese, L. J. Magrum y G. E. Fox en *Journal of Molecular Evolution*, vol. 11, n.° 3, págs. 245-252; 1978.

METHANOGENS: REEVALUATION OF A UNIQUE BIOLOGICAL GROUP. W. E. Balch, G. E. Fox, L. J. Magrum, C. R. Woese y R. S. Wolfe en *Microbiological Reviews*, vol. 43, n.° 2, págs. 260-296; junio, 1979.

RESPIRATION IN METHANOGENIC BACTERIA. R. S. Wolfe en *Diversity of Bacterial Respiratory Systems*, dirigido por C. R. Knowles. CRC Press, Inc., 1980.

THE PHYLOGENY OF PROKARYOTES. G. E. Fox, E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen y C. R. Woese en *Science*, vol. 209, n.° 4455, págs. 457-463; 25 de julio de 1980.

ASIGNACION DE RECURSOS POR PROGRAMACION LINEAL

LINEAR PROGRAMMING AND ECONOMIC ANALYSIS. Robert Dorfman, Paul A. Samuelson y Robert M. Solow. McGraw-Hill Book Company, 1958.

THE THEORY OF LINEAR ECONOMIC MODELS. David Gale. McGraw-Hill Book Company, 1960.

LINEAR PROGRAMMING AND EXTENSIONS. G. B. Dantzig. Princeton University Press, 1965.

APPLIED MATHEMATICAL PROGRAMMING. Stephen P. Bradley, Arnoldo C. Hax y Thomas L. Magnanti. Addison-Wesley Publishing Company, 1977.

A COURSE IN LINEAR PROGRAMMING. V. Chvátal. W. H. Freeman and Company, en prensa.

REGULACION DE LA TEMPERATURA EN EL ENJAMBRE DE ABEJAS MELIFERAS

THE SOCIAL BEHAVIOR OF BEES: A COMPARATIVE STUDY. Charles D. Michener. Harvard University Press, 1974.

THE NATURAL HISTORY OF THE FLIGHT OF HONEYBEE SWARMS. T. D. Seeley,

FORMACION DE LA TIERRA A PARTIR DE PLANETESIMALES

ORIGIN OF THE EARTH AND MOON. A. E. Ringwood. Springer-Verlag, 1979.

FORMATION OF THE TERRESTRIAL PLANETS. George W. Wetherill en *Annual Review of Astronomy and Astrophysics*, vol. 18, págs. 77-113; 1980.

LOS "ICTINEOS" DE NARCIS MONTURIOL

ENSAYO SOBRE EL ARTE DE NAVEGAR POR DEBAJO DEL AGUA. Narcís Monturiol. (Prólogo de Joan Mañé i Flaguer y un escrito de Josep Pascual i Deop.) Imprenta de Henrich C.ª en comandita. Sucesores de N. Ramírez y compañía. Barcelona, 1891.

NARCISO MONTURIOL Y LA NAVEGACIÓN SUBMARINA. JUICIOS CRÍTICOS EMITIDOS SOBRE LOS IMPORTANTÍSIMOS TRABAJOS REALIZADOS POR ESTE SABIO INVENTOR CATALÁN, COLECCIONADOS POR EL DR. D. JERÓNIMO ESTRANY. G. Gili, Barcelona 1915.

VIDA D'HEROI. NARCÍS MONTURIOL: INVENTOR DE LA NAVEGACIÓ SUBMARINA. Josep Puig i Pujadas. Imprenta L'avenç. Barcelona, 1918.

JUEGOS MATEMATICOS

SYMMETRY. Hermann Weyl. Princeton University Press, 1952.

ROTATIONS AND REFLECTIONS. Martin Gardner en *The Unexpected Hanging and Other Matematical Diversions*. Simon y Schuster, 1969.

SYMMETRY IN SCIENCE AND ART. A. V. Schubnikov y V. A. Koptsik. Plenum Press, 1974.

TALLER Y LABORATORIO

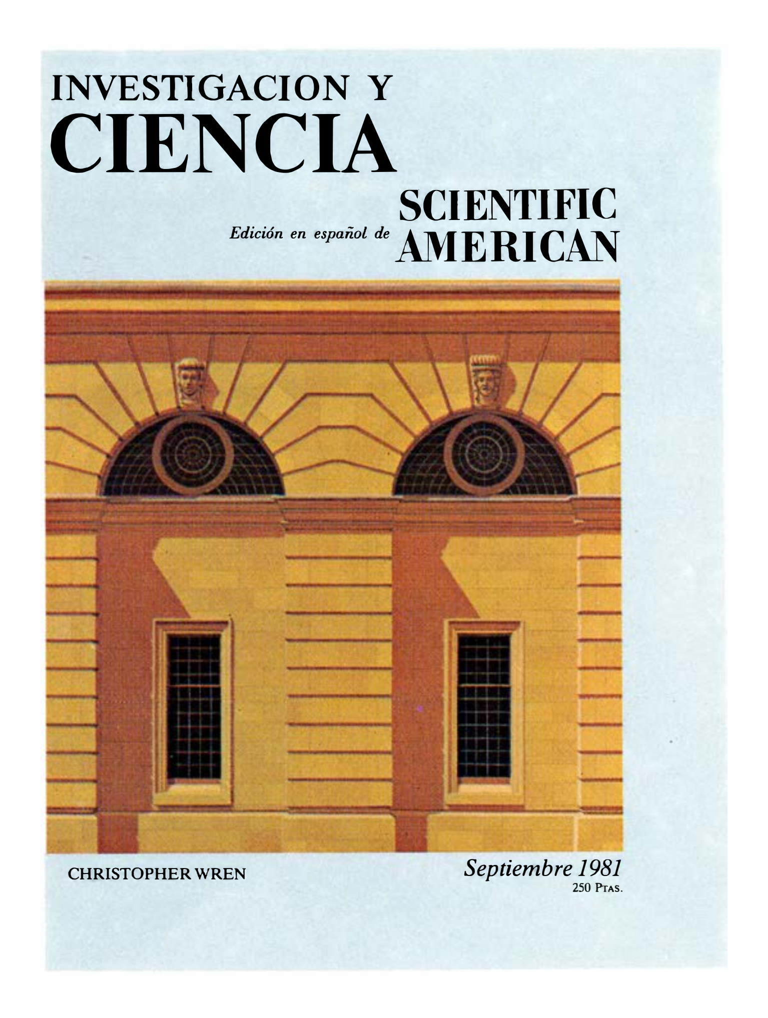
THE FUNCTIONAL PROPERTIES OF THE EGG WHITE PROTEINS. L. R. MacDonell, R. E. Feeney, H. L. Hanson, Agnes Campbell y T. F. Sugihara en *Food Technology*, vol. 9, n.° 2, págs. 49-53; febrero, 1955.

PROTEINS. Vicki Cobb en *Science Experiments You can Eat*, J. B. Lippincott Company, 1972.

PROTEÍNAS: CONFORMACIÓN TRIDIMENSIONAL. Albert L. Lehninger en *Bioquímica*. Editorial Omega, 1977.

MIRACLES IN A SHELL (EGG COOKERY). Madeleine Kamman en *The Making of a Cook*. Atheneum Publishers, 1978.

Seguiremos explorando los campos del conocimiento



EL POTASIO Y EL SODIO EN LAS CELULAS VIVAS, por A. Rodríguez Navarro
El potasio es el catión más abundante en las células vivas, aunque el sodio es el catión más abundante en los biotopos. Ambos cationes están implicados en complejos mecanismos de transporte.

LA SALINIDAD DE LOS RIOS, por Arthur F. Pillsbury
Los ríos vierten al océano las sales que se disuelven de las rocas. Sin embargo, al retenerse el caudal para el riego, la evapotranspiración concentra las sales en el suelo.

LA ATMOSFERA DE VENUS, por Gerald Schubert y Curt Covey
Diez años de exploración aerospacial revelan que se compone, casi por entero, de dióxido de carbono. Sus nubes de ácido sulfúrico se ven arrastradas por vientos de hasta 360 kilómetros por hora.

INGENIERIA GENETICA EN CELULAS DE MAMIFERO, por W. French Anderson y Elaine G. Diacumakos
Uno de los fines de la tecnología del ADN recombinante es la curación de enfermedades genéticas humanas. Se ha dado un paso adelante hacia ese objetivo: la corrección de un defecto hereditario.

HACES DE FIBRAS Y TEORIA CUANTICA, por Herbert J. Bernstein y Anthony V. Phillips
Una rama de la matemática, que extiende la noción de curvatura a análogos topológicos de una banda de Möbius, puede explicar las actuales teorías sobre interacciones de partículas elementales.

MIMETISMO EN LAS SEÑALES SEXUALES DE LAS LUCIERNAGAS, por James E. Lloyd
Los machos de las luciérnagas emiten destellos mientras vuelan en busca de pareja. Las hembras de su especie les responden desde el suelo. Las hembras de otras especies también emiten respuestas.

LA ARQUITECTURA DE CHRISTOPHER WREN, por Harold Dorn y Robert Mark
Fue un destacado arquitecto y eminente hombre de ciencia, que no utilizó jamás en los proyectos de sus edificios los principios de la mecánica teórica, que conocía.

VALORES Y ACTITUDES DEL PUEBLO POLACO, por Stefan Nowak
Los recientes sucesos de Polonia, a la luz de 25 años de investigación social, no parecen derivarse de un cambio de valores, sino de la exigencia de instituciones sociales más de acuerdo con esos valores, mantenidos con tesón.

